

The \mathbb{Z}_2 network model for the quantum spin Hall effect: two-dimensional Dirac fermions, topological quantum numbers, and corner multifractality

Shinsei Ryu¹, Christopher Mudry², Hideaki Obuse³ and Akira Furusaki⁴

¹Department of Physics, University of California, Berkeley, CA 94720, USA

²Condensed Matter Theory Group, Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland

³Department of Physics, Kyoto University, Kyoto 606-8502, Japan

⁴Condensed Matter Theory Laboratory, RIKEN, Wako, Saitama 351-0198, Japan

E-mail: sryu@berkeley.edu, christopher.mudry@psi.ch,
obuse@scphys.kyoto-u.ac.jp, furusaki@riken.jp

Abstract. The quantum spin Hall effect shares many similarities (and some important differences) with the quantum Hall effect for the electric charge. As with the quantum (electric charge) Hall effect, there exists a correspondence between bulk and boundary physics that allows to characterize the quantum spin Hall effect in diverse and complementary ways. In this paper, we derive from the network model that encodes the quantum spin Hall effect, the so-called \mathbb{Z}_2 network model, a Dirac Hamiltonian in two dimensions. In the clean limit of this Dirac Hamiltonian, we show that the bulk Kane-Mele \mathbb{Z}_2 invariant is nothing but the $SU(2)$ Wilson loop constructed from the $SU(2)$ Berry connection of the occupied Dirac-Bloch single-particle states. In the presence of disorder, the non-linear sigma model (NLSM) that is derived from this Dirac Hamiltonian describes a metal-insulator transition in the standard two-dimensional symplectic universality class. In particular, we show that the fermion doubling prevents the presence of a topological term in the NLSM that would change the universality class of the ordinary two-dimensional symplectic metal-insulator transition. This analytical result is fully consistent with our previous numerical studies of the bulk critical exponents at the metal-insulator transition encoded by the \mathbb{Z}_2 network model. Finally, we improve the quality and extend the numerical study of boundary multifractality in the \mathbb{Z}_2 topological insulator. We show that the hypothesis of two-dimensional conformal invariance at the metal-insulator transition is verified within the accuracy of our numerical results.

PACS numbers: 73.20.Fz, 71.70.Ej, 73.43.-f, 05.45.Df

Contents

1	Introduction	2
2	Definition of the \mathbb{Z}_2 network model for the quantum spin Hall effect	5
3	Two-dimensional Dirac Hamiltonian from the \mathbb{Z}_2 network model	8
3.1	Change of the basis for the scattering states and one-step time evolution	8
3.2	Dirac Hamiltonian close to $\theta = 0$	10
3.3	\mathbb{Z}_2 topological number	13
4	Numerical study of boundary multifractality in the \mathbb{Z}_2 network model	23
4.1	Boundary and corner multifractality	24
4.2	Boundary condition changing operator	28
5	Conclusions	30

1. Introduction

Spin-orbit coupling has long been known to be essential to account for the band structure of semiconductors, say, semiconductors with the zinc-blende crystalline structure. Monographs have been dedicated to reviewing the effects of the spin-orbit coupling on the Bloch bands of conductors and semiconductors [1]. Electronic transport properties of metals and semiconductors in which impurities are coupled to the conduction electrons by the spin-orbit coupling, i.e., when the impurities preserve the time-reversal symmetry but break the spin-rotation symmetry, are also well understood since the prediction of weak antilocalization effects [2]. Hence, the prediction of the quantum spin Hall effect in two-dimensional semiconductors with time-reversal symmetry but a sufficiently strong breaking of spin-rotation symmetry is rather remarkable in view of the maturity of the field dedicated to the physics of semiconductors [3, 4, 5, 6]. The quantum spin Hall effect was observed in HgTe/(Hg,Cd)Te quantum wells two years later [7]. Even more remarkably, this rapid progress was followed by the prediction of three-dimensional topological insulators [8, 9, 10] and its experimental confirmation for Bi-based compounds [11, 12, 13, 14, 15].

The quantum spin Hall effect, like its relative, the quantum (electric charge) Hall effect, can be understood either as a property of the two-dimensional bulk or as a property of the one-dimensional boundary. The bulk can be characterized by certain integrals over the Brillouin zone of Berry connections calculated from Bloch eigenstates. These integrals are only allowed to take discrete values and are examples of topological invariants from the mathematical literature. As is well known, the topological number ν takes integer values for the quantum (electric charge) Hall effect [16]. By contrast, it takes only two distinct values ($\nu = 0$ or 1) for time-reversal invariant, \mathbb{Z}_2 topological band insulators [4, 8, 9, 10, 17]. Because they are quantized, they cannot change under a

small continuous deformation of the Hamiltonian, including a perturbation that breaks translation invariance, i.e., disorder.

The bulk topological quantum numbers are closely connected with the existence of stable gapless edge states along the boundary of a topological insulator, or more precisely along the interface between two insulators with different topological numbers. The number of gapless edge modes is determined by the difference of the topological numbers. On the edge of a two-dimensional \mathbb{Z}_2 topological band insulator with $\nu = 1$, there exists helical edge states, a Kramers' pair of counter propagating modes, which interpolates between the bulk valence band and the bulk conduction band. If one changes the Fermi energy from the center of the band gap to lower energies through the conduction band, one should observe a transition from a \mathbb{Z}_2 topological insulator to a metal, and then from a metal to a trivial band insulator ($\nu = 0$) without helical edge states. Since both helical edge states and a metallic phase are stable against (weak) disorder (due to the quantized topological number and to weak anti-localization, respectively), the same sequence of phases should appear as the Fermi energy is varied even in the presence of disorder, as confirmed recently by numerical simulations [18, 19]. A question one can naturally ask is then whether there is any difference between the critical phenomena at the metal-to- \mathbb{Z}_2 -topological-insulator transition and those at the metal-to-trivial-insulator transition. This is the question which we revisit in this paper, extending our previous studies [19, 20]. It will become clear that one needs to distinguish between bulk and boundary properties in the universal critical phenomena.

For the quantum (electric charge) Hall effect, the Chalker-Coddington network model serves as a standard model for studying critical properties at Anderson transition between different quantum Hall states [21]. The elementary object in the Chalker-Coddington network model is chiral edge states. These edge states are plane waves propagating along the links of each plaquette which represents a puddle of a quantum Hall droplet formed in the presence of spatially slowly varying potential. They are chiral as they represent the mode propagating along equipotential lines in the direction determined by the external magnetic field. The Chalker-Coddington network model is a unitary scattering matrix that scales in size with the number of links defining the network, and with a deterministic parameter that quantifies the relative probability for an incoming mode to scatter into a link rotated by $+\pi/2$ or $-\pi/2$. By tuning this parameter through the value $1/2$, one can go through a transition from one insulating phase to another insulating phase, with the topological number ν changed by one. This remains true even when the phase of an edge state along any link is taken to be an independent random number to mimic the effects of static local disorder. The Chalker-Coddington model is a powerful tool to characterize the effects of static disorder on the direct transition between two successive integer quantum Hall states. It has demonstrated that this transition is continuous and several critical exponents at this transition have been measured from the Chalker-Coddington model [21, 22].

The present authors have constructed in [19] a generalization of the Chalker-Coddington model that describes the physics of the two-dimensional quantum spin Hall

effect. We shall call this network model the \mathbb{Z}_2 network model, which will be briefly reviewed in section 2. As with the Chalker-Coddington model, edge states propagate along the links of each plaquette of the square lattice. Unlike the Chalker-Coddington model there are two edge states per link that form a single Kramers' doublet, which corresponds to helical edge states moving along a puddle of a quantum spin Hall droplet. Kramers' doublets undergo the most general unitary scattering compatible with time-reversal symmetry at the nodes of the square lattice. The \mathbb{Z}_2 network model is thus a unitary scattering matrix that scales in size with the number of links defining the network and that preserves time-reversal symmetry. The \mathbb{Z}_2 network model supports one metallic phase and two insulating phases, as we discussed earlier[‡]. The metallic phase prevents any direct transition between the insulating phases and the continuous phase transition between the metallic and any of the insulating phases belongs to the two-dimensional symplectic universality class of Anderson localization [2].

Numerical simulations have shown that bulk properties at metal-insulator transition in the \mathbb{Z}_2 network model are the same as those at conventional metal-insulator transitions in the two-dimensional symplectic symmetry class [19, 20]. In fact, one can understand this result from the following general argument based on universality. The non-linear sigma model (NLSM) description is a very powerful, standard theoretical approach to Anderson metal-insulator transition [23]. A NLSM can have a topological term if the homotopy group of the target manifold, which is determined by the symmetry of the system at hand, is nontrivial. Interestingly, in the case of the symplectic symmetry class, as is called the statistical ensemble of systems (including quantum spin Hall systems) that are invariant under time reversal but are not invariant under $SU(2)$ spin rotation, the NLSM admits a \mathbb{Z}_2 topological term [24, 25, 26]. Moreover, the NLSM in the symplectic symmetry class with a \mathbb{Z}_2 topological term cannot support an insulating phase. This can be seen from the fact that this NLSM describes surface Dirac fermions of a three-dimensional \mathbb{Z}_2 topological insulator which are topologically protected from Anderson localization [27, 28, 29]. This in turn implies that any two-dimensional metal-insulator transition in time-reversal-invariant but spin-rotation-noninvariant systems should be in the same and unique universality class that is encoded by the NLSM without a topological term in the (ordinary) symplectic class.

Whereas bulk critical properties at the transition between a metal and a \mathbb{Z}_2 topological insulator do not depend on the topological nature of the insulating phase, there are boundary properties that can distinguish between a topologically trivial and non-trivial insulating phases. Boundary multifractality is a very convenient tool to probe any discrepancy between universal bulk and boundary properties at Anderson transition [30, 31]. To probe this difference, the present authors performed a multifractal analysis of the edge states that propagate from one end to the other in a network model at criticality with open boundary condition in the transverse direction [20]. It was found that boundary multifractal exponents are sensitive to the presence or absence of a helical

[‡] The presence or absence of a single helical edge state in an insulating phase is solely dependent on the boundary conditions which one imposes on the network model.

Kramers' doublet propagating along the boundary.

The goal of this paper is 2-fold:

- (i) to establish a direct connection between the \mathbb{Z}_2 network model and a Hamiltonian description of the \mathbb{Z}_2 topological insulator perturbed by time-reversal symmetric local static disorder.
- (ii) to improve the quality and extend the numerical study of boundary multifractality in the \mathbb{Z}_2 topological insulator.

For item (i), in section 3, we are going to relate the \mathbb{Z}_2 network model to a problem of Anderson localization in the two-dimensional symplectic universality class that is encoded by a stationary 4×4 Dirac Hamiltonian perturbed by static disorder that preserves time-reversal symmetry but breaks spin-rotation symmetry. This result is a natural generalization of the fact that the Chalker-Coddington network model can be related [32] to a 2×2 Dirac Hamiltonian with static disorder [33]. In the clean limit, we shall characterize the \mathbb{Z}_2 insulating phases in the 4×4 Dirac Hamiltonian by a \mathbb{Z}_2 topological invariant. In particular, we show that an $SU(2)$ Wilson loop of Berry connection of Bloch wave functions is equivalent to the \mathbb{Z}_2 index introduced by Kane and Mele [4]. The 4×4 Dirac Hamiltonian will allow us to make contact between the \mathbb{Z}_2 network model and the NLSM description of two-dimensional Anderson localization in the symplectic universality class derived 30 years ago by Hikami et al. in [2]. In our opinion, this should remove any lingering doubts that the metal-insulator transition between a two-dimensional metallic state and a two-dimensional \mathbb{Z}_2 insulator that is driven by static disorder is anything but conventional.

For item (ii), besides improving the accuracy of the critical exponents for one-dimensional boundary multifractality in the \mathbb{Z}_2 network model, we compute critical exponents for two zero-dimensional boundaries (corners) in section 4. We shall use these critical exponents to verify the hypothesis that conformal invariance holds at the metal-insulator transition and imposes relations between lower-dimensional boundary critical exponents.

2. Definition of the \mathbb{Z}_2 network model for the quantum spin Hall effect

The \mathbb{Z}_2 network model is defined as follows. First, one draws a set of corner sharing square plaquettes on the two-dimensional Cartesian plane. Each edge of a plaquette is assigned two opposite directed links. This is the network. There are two types S and S' of shared corners, which we shall call the nodes of the network. Second, we assign to each directed link an amplitude ψ , i.e., a complex number $\psi \in \mathbb{C}$. Any amplitude ψ is either an incoming or outgoing plane wave that undergoes a unitary scattering process at a node. We also assign a 4×4 unitary matrix S to each node of the network. The set of all directed links obeying the condition that they are either the incoming or outgoing plane waves of the set of all nodal unitary scattering matrices defines a solution to the \mathbb{Z}_2 network model.

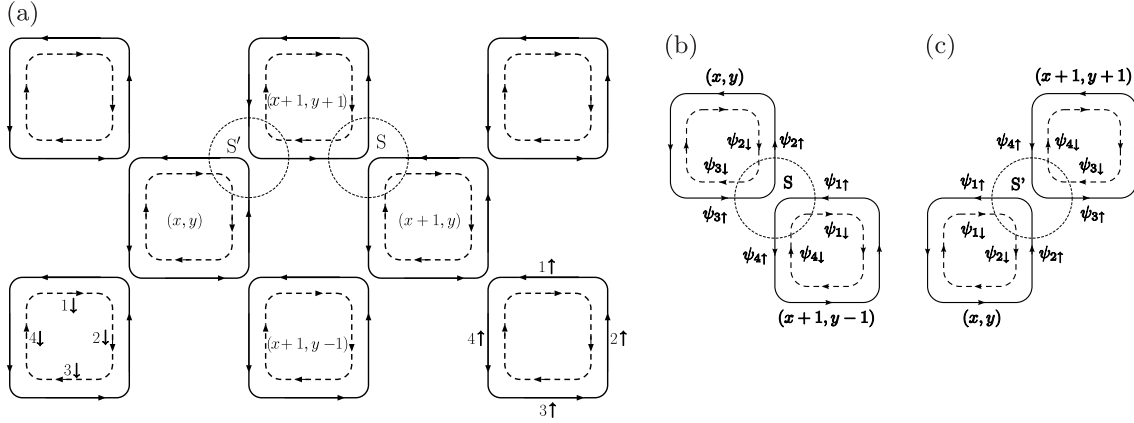


Figure 1. (a) The \mathbb{Z}_2 network model. The solid and dashed lines represent the links for up and down spin electrons, respectively. The electrons are unitarily scattered at the nodes S and S' . The choice for the scattering basis at the nodes S and S' is shown in (b) and (c), respectively.

To construct an explicit representation of the \mathbb{Z}_2 network model, the center of each plaquette is assigned the coordinate (x, y) with x and y taking integer values, as is done in figure 1. We then label the 8 directed links $\psi_{n\sigma}(x, y)$ of any given plaquette by the coordinate (x, y) of the plaquette, the side $n = 1, 2, 3, 4$ of the plaquette with the convention shown in figure 1, and the spin index $\sigma = \uparrow$ or $\sigma = \downarrow$ if the link is directed counterclockwise or clockwise, respectively, relative to the center of the plaquette. The 4×4 unitary S -matrix is then given by

$$\begin{pmatrix} \psi_{2\uparrow}(x, y) \\ \psi_{3\downarrow}(x, y) \\ \psi_{4\uparrow}(x+1, y-1) \\ \psi_{1\downarrow}(x+1, y-1) \end{pmatrix} =: S \begin{pmatrix} \psi_{3\uparrow}(x, y) \\ \psi_{2\downarrow}(x, y) \\ \psi_{1\uparrow}(x+1, y-1) \\ \psi_{4\downarrow}(x+1, y-1) \end{pmatrix} \quad (1)$$

at any node of type S or as

$$\begin{pmatrix} \psi_{3\uparrow}(x+1, y+1) \\ \psi_{4\downarrow}(x+1, y+1) \\ \psi_{1\uparrow}(x, y) \\ \psi_{2\downarrow}(x, y) \end{pmatrix} =: S' \begin{pmatrix} \psi_{4\uparrow}(x+1, y+1) \\ \psi_{3\downarrow}(x+1, y+1) \\ \psi_{2\uparrow}(x, y) \\ \psi_{1\downarrow}(x, y) \end{pmatrix} \quad (2)$$

at any node of type S' , with

$$S = U(x, y)S_0V(x, y), \quad S' = U'(x, y)S_0V'(x, y). \quad (3)$$

Here, the 4×4 unitary matrix

$$S_0 := \begin{pmatrix} rs_0 & tQ \\ -tQ^\dagger & rs_0 \end{pmatrix} \quad (4)$$

is presented with the help of the unit 2×2 matrix s_0 and of the 2×2 matrix

$$Q := s_1 \sin \theta + s_3 \cos \theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}, \quad (5)$$

(s_1 , s_2 , and s_3 are the 2×2 Pauli matrices) that are both acting on the spin indices $\sigma = \uparrow, \downarrow$, together with the real-valued parameters

$$r := \tanh X, \quad t := \frac{1}{\cosh X}, \quad (6)$$

with

$$\{(X, \theta) \mid 0 \leq X \leq \infty, \quad 0 \leq \theta \leq \pi/2\}. \quad (7)$$

For later use, we shall also introduce the real-valued parameter $\beta \in [0, \pi]$ through

$$r = \cos \beta, \quad t = \sin \beta. \quad (8)$$

The parameter θ controls the probability of spin-flip scattering, $\sin^2 \theta$. The unitary matrices U, V, U', V' are defined as

$$U(x, y) = \text{diag}(e^{i\chi_2(x, y)}, e^{i\chi_3(x, y)}, e^{i\chi_4(x+1, y-1)}, e^{i\chi_1(x+1, y-1)}), \quad (9a)$$

$$V(x, y) = \text{diag}(e^{i\chi_3(x, y)}, e^{i\chi_2(x, y)}, e^{i\chi_1(x+1, y-1)}, e^{i\chi_4(x+1, y-1)}), \quad (9b)$$

$$U'(x, y) = \text{diag}(e^{i\chi_3(x+1, y+1)}, e^{i\chi_4(x+1, y+1)}, e^{i\chi_1(x, y)}, e^{i\chi_2(x, y)}), \quad (9c)$$

$$V'(x, y) = \text{diag}(e^{i\chi_4(x+1, y+1)}, e^{i\chi_3(x+1, y+1)}, e^{i\chi_2(x, y)}, e^{i\chi_1(x, y)}), \quad (9d)$$

where $2\chi_n(x, y)$ equals a (random) phase that wave functions acquire when propagating along the edge n of the plaquette centered at (x, y) .

The \mathbb{Z}_2 network model is uniquely defined from the scattering matrices S and S' . By construction, the S -matrix is time-reversal symmetric, i.e.,

$$\begin{pmatrix} is_2 & 0 \\ 0 & is_2 \end{pmatrix} S^* \begin{pmatrix} -is_2 & 0 \\ 0 & -is_2 \end{pmatrix} = S^\dagger, \quad (10)$$

and a similar relation holds for S' .

In [19], we obtained the phase diagram of the \mathbb{Z}_2 network model shown schematically in figure 2(a). Thereto, (X, θ) are spatially uniform deterministic parameters that can be changed continuously. On the other hand, the phases χ_n of all link plane waves in the \mathbb{Z}_2 network model are taken to be independently and uniformly distributed random variables over the range $[0, 2\pi)$. The line $\theta = 0$ is special in that the \mathbb{Z}_2 network model reduces to two decoupled Chalker-Coddington network models [19]. Along the line $\theta = 0$, the point

$$X_{\text{CC}} = \ln(1 + \sqrt{2}) \iff \beta = \frac{\pi}{4} \quad (11)$$

realizes a quantum critical point that separates two insulating phases differing by one gapless edge state or, equivalently, by one unit in the Hall conductivity, per spin. Alternatively, θ can also be chosen to be randomly and independently distributed at each node with the probability $\sin(2\theta)$ over the range $(0, \pi/2)$. This leaves X as the sole deterministic parameter that controls the phase diagram as shown in figure 2(b). When performing numerically a scaling analysis with the size of the \mathbb{Z}_2 network model, one must account for the deviations away from one-parameter scaling induced by irrelevant operators. The \mathbb{Z}_2 network model with a randomly distributed θ minimizes such finite-size effects (see [19]).

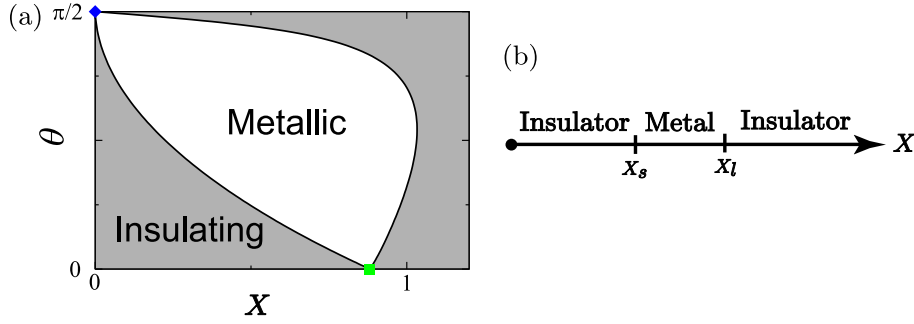


Figure 2. (a) Schematic phase diagram from the analysis of the \mathbb{Z}_2 network model with the constant X and θ . The metallic phase is surrounded by the two insulating phases with the critical points X_s and $X_l(> X_s)$ for $0 < \theta < \pi/2$. The fixed point denoted by a filled (green) square along the boundary $\theta = 0$ is the unstable quantum critical point located at $X_{CC} = \ln(1 + \sqrt{2})$ separating two insulating phases in the Chalker-Coddington model. The fixed point denoted by the filled (blue) rhombus at the upper left corner is the unstable metallic phase. The shape of the metallic phase is controlled by the symmetry crossover between the unitary and symplectic symmetry classes. (b) The phase diagram for \mathbb{Z}_2 network model with randomly distributed θ over the range $(0, \pi/2)$.

3. Two-dimensional Dirac Hamiltonian from the \mathbb{Z}_2 network model

The Chalker-Coddington model is related to the two-dimensional Dirac Hamiltonian as was shown by Ho and Chalker in [32]. We are going to establish the counterpart of this connection for the \mathbb{Z}_2 network model. A unitary matrix is the exponential of a Hermitian matrix. Hence, our strategy to construct a Hamiltonian from the \mathbb{Z}_2 network model is going to be to view the unitary scattering matrix of the \mathbb{Z}_2 network model as a unitary time evolution whose infinitesimal generator is the sought Hamiltonian. To this end, we proceed in two steps in order to present the \mathbb{Z}_2 network model into a form in which it is readily interpreted as a unitary time evolution. First, we change the choice of the basis for the scattering states and select the proper unit of time. We then perform a continuum approximation, by which the \mathbb{Z}_2 network model is linearized, so to say. This will yield an irreducible 4-dimensional representation of the Dirac Hamiltonian in $(2+1)$ -dimensional space and time, a signature of the fermion doubling when deriving a continuum Dirac Hamiltonian from a time-reversal symmetric and local two-dimensional lattice model.

3.1. Change of the basis for the scattering states and one-step time evolution

Our goal is to reformulate the \mathbb{Z}_2 network model defined in Sec. 2 in such a way that the scattering matrix maps incoming states into outgoing states sharing the *same* internal and space labels but a different “time” label. This involves a change of basis for the scattering states and an “enlargement” of the Hilbert space spanned by the scattering states. The parameter θ is assumed to be spatially uniform. We choose the plaquette

(x, y) of the network.

At node **S** of the plaquette (x, y) , we make the basis transformation and write the S -matrix (1) in the form

$$\begin{pmatrix} \psi_{1\downarrow} \\ \psi_{3\downarrow} \\ \psi_{2\uparrow} \\ \psi_{4\uparrow} \end{pmatrix} =: \mathcal{M}_S \begin{pmatrix} \psi_{1\uparrow} \\ \psi_{3\uparrow} \\ \psi_{2\downarrow} \\ \psi_{4\downarrow} \end{pmatrix}, \quad \mathcal{M}_S = \mathcal{U} \mathcal{N}_S \mathcal{U}, \quad (12)$$

where we have defined

$$\mathcal{N}_S = \begin{pmatrix} 0 & -t t_-^x t_+^y \sin \theta & t t_-^x t_+^y \cos \theta & r \\ t t_+^x t_-^y \sin \theta & 0 & r & -t t_+^x t_-^y \cos \theta \\ t t_+^x t_-^y \cos \theta & r & 0 & t t_+^x t_-^y \sin \theta \\ r & -t t_-^x t_+^y \cos \theta & -t t_-^x t_+^y \sin \theta & 0 \end{pmatrix} \quad (13)$$

and

$$\mathcal{U}(x, y) = \text{diag}(e^{i\chi_1(x, y)}, e^{i\chi_3(x, y)}, e^{i\chi_2(x, y)}, e^{i\chi_4(x, y)}). \quad (14)$$

Here given $n = 1, 2, 3, 4$ and $\sigma = \uparrow, \downarrow$, we have introduced the shift operators acting on $\psi_{n\sigma}(x, y)$,

$$t_{\pm}^x \psi_{n\sigma}(x, y) := \psi_{n\sigma}(x \pm 1, y), \quad t_{\pm}^y \psi_{n\sigma}(x, y) := \psi_{n\sigma}(x, y \pm 1), \quad (15)$$

and similarly on the phases $\chi_n(x, y) \in [0, 2\pi)$. We note that the scattering matrix \mathcal{N}_S is multiplied by the unitary matrix \mathcal{U} from the left and the right in (12), because the Kramers' doublet acquires exactly the same phase χ_n when traversing on the edge n of the plaquette (x, y) before and after experiencing the scattering \mathcal{N}_S at the node **S**.

At node **S'** of the plaquette (x, y) , we make the basis transformation and rewrite the scattering matrix S' (2) into the form

$$\begin{pmatrix} \psi_{1\uparrow} \\ \psi_{3\uparrow} \\ \psi_{2\downarrow} \\ \psi_{4\downarrow} \end{pmatrix} =: \mathcal{M}_{S'} \begin{pmatrix} \psi_{1\downarrow} \\ \psi_{3\downarrow} \\ \psi_{2\uparrow} \\ \psi_{4\uparrow} \end{pmatrix}, \quad \mathcal{M}_{S'} = \mathcal{U} \mathcal{N}_{S'} \mathcal{U}, \quad (16)$$

where we have defined

$$\mathcal{N}_{S'} = \begin{pmatrix} 0 & -t t_+^x t_+^y \sin \theta & r & -t t_+^x t_+^y \cos \theta \\ t t_-^x t_-^y \sin \theta & 0 & t t_-^x t_-^y \cos \theta & r \\ r & t t_+^x t_+^y \cos \theta & 0 & -t t_+^x t_+^y \sin \theta \\ -t t_-^x t_-^y \cos \theta & r & t t_-^x t_-^y \sin \theta & 0 \end{pmatrix}. \quad (17)$$

As it should be

$$\mathcal{M}_S^\dagger \mathcal{M}_S = \mathcal{M}_{S'}^\dagger \mathcal{M}_{S'} = 1. \quad (18)$$

Next, we introduce the discrete time variable $l \in \mathbb{Z}$ as follows. We define the elementary discrete unitary time evolution to be

$$\begin{pmatrix} \psi_{+\downarrow} \\ \psi_{-\uparrow} \\ \psi_{+\uparrow} \\ \psi_{-\downarrow} \end{pmatrix}_{l+1} := \begin{pmatrix} 0 & \mathcal{M}_S \\ \mathcal{M}_{S'} & 0 \end{pmatrix} \begin{pmatrix} \psi_{+\downarrow} \\ \psi_{-\uparrow} \\ \psi_{+\uparrow} \\ \psi_{-\downarrow} \end{pmatrix}_l. \quad (19)$$

Here, to treat on equal footing the nodes of type S and S' , we have enlarged the scattering basis with the introduction of the doublets

$$\psi_+ := \begin{pmatrix} \psi_1 \\ \psi_3 \end{pmatrix}, \quad \psi_- := \begin{pmatrix} \psi_2 \\ \psi_4 \end{pmatrix}. \quad (20)$$

Due to the off-diagonal block structure in the elementary time evolution, it is more convenient to consider the “one-step” time evolution operator defined by

$$\begin{aligned} \begin{pmatrix} \psi_{+\downarrow} \\ \psi_{-\uparrow} \\ \psi_{+\uparrow} \\ \psi_{-\downarrow} \end{pmatrix}_{l+2} &= \begin{pmatrix} \mathcal{M}_S \mathcal{M}_{S'} & 0 \\ 0 & \mathcal{M}_{S'} \mathcal{M}_S \end{pmatrix} \begin{pmatrix} \psi_{+\downarrow} \\ \psi_{-\uparrow} \\ \psi_{+\uparrow} \\ \psi_{-\downarrow} \end{pmatrix}_l \\ &\equiv \begin{pmatrix} \mathcal{M}_{SS'} & 0 \\ 0 & \mathcal{M}_{S'S} \end{pmatrix} \begin{pmatrix} \psi_{+\downarrow} \\ \psi_{-\uparrow} \\ \psi_{+\uparrow} \\ \psi_{-\downarrow} \end{pmatrix}_l. \end{aligned} \quad (21)$$

The two Hamiltonians generating this unitary time evolution are then

$$\mathcal{H}_{SS'} := +i \ln \mathcal{M}_{SS'}, \quad \mathcal{H}_{S'S} := +i \ln \mathcal{M}_{S'S}. \quad (22)$$

Evidently, the additivity of the logarithm of a product implies that

$$\mathcal{H}_{SS'} = \mathcal{H}_{S'S}. \quad (23)$$

From now on, we will consider $\mathcal{H}_{SS'}$ exclusively since $\mathcal{M}_{S'S} = \exp(i\mathcal{H}_{S'S})$ merely duplicates the information contains in $\mathcal{M}_{SS'} = \exp(i\mathcal{H}_{SS'})$.

3.2. Dirac Hamiltonian close to $\theta = 0$

In this section, we are going to extract from the unitary time-evolution (21)–(23) of the \mathbb{Z}_2 network model a 4×4 continuum Dirac Hamiltonian in the close vicinity of the quantum critical point

$$(\theta, \beta)_{CC} := (0, \pi/4). \quad (24)$$

To this end and following [32], it is convenient to measure the link phases χ_n ($n = 1, 2, 3, 4$) relative to their values when they carry a flux of π per plaquette. Hence, we redefine

$$\chi_4 \rightarrow \chi_4 + \frac{\pi}{2} \quad (25)$$

on all plaquettes.

Our strategy consists in performing an expansion of

$$\mathcal{H}_{SS'} = +i \ln \mathcal{M}_{SS'} = +i (\ln \mathcal{M}_S + \ln \mathcal{M}_{S'}) = +i \ln \mathcal{M}_{S'S} = \mathcal{H}_{S'S} \quad (26)$$

defined in (22) to leading order in powers of

$$\theta, \quad \frac{m}{2} \equiv \beta - \frac{\pi}{4}, \quad \partial_{x,y} \equiv \ln t_+^{x,y}, \quad \chi_n \quad (27)$$

with $n = 1, 2, 3, 4$ where $\partial_{x,y}$ is the generator of infinitesimal translation on the network (the two-dimensional momentum operator).

When $\theta = 0$, the unitary time-evolution operator at the plaquette (x, y) is given by

$$\begin{pmatrix} \psi_{+\downarrow} \\ \psi_{-\uparrow} \end{pmatrix}_{l+2} = \mathcal{M}_{SS'}^{(0)} \begin{pmatrix} \psi_{+\downarrow} \\ \psi_{-\uparrow} \end{pmatrix}_l, \quad (28)$$

$$\mathcal{M}_{SS'}^{(0)} = \begin{pmatrix} A^{(0)} D^{(0)} & 0 \\ 0 & B^{(0)} C^{(0)} \end{pmatrix}, \quad (29)$$

whereby

$$\mathcal{M}_{SS'}^{(0)} = \mathcal{M}_S^{(0)} \mathcal{M}_{S'}^{(0)}, \quad (30)$$

$$\mathcal{M}_S^{(0)} = \begin{pmatrix} 0 & A^{(0)} \\ B^{(0)} & 0 \end{pmatrix}, \quad \mathcal{M}_{S'}^{(0)} = \begin{pmatrix} 0 & C^{(0)} \\ D^{(0)} & 0 \end{pmatrix}, \quad (31)$$

with the 2×2 operator-valued matrices

$$A^{(0)} := \begin{pmatrix} e^{i\chi_1} t_-^x t_+^y e^{i\chi_2} \sin \beta & i e^{i(\chi_1 + \chi_4)} \cos \beta \\ e^{i(\chi_3 + \chi_2)} \cos \beta & -i e^{i\chi_3} t_+^x t_-^y e^{i\chi_4} \sin \beta \end{pmatrix}, \quad (32)$$

$$B^{(0)} := \begin{pmatrix} e^{i\chi_2} t_+^x t_-^y e^{i\chi_1} \sin \beta & e^{i(\chi_2 + \chi_3)} \cos \beta \\ i e^{i(\chi_4 + \chi_1)} \cos \beta & -i e^{i\chi_4} t_-^x t_+^y e^{i\chi_3} \sin \beta \end{pmatrix}, \quad (33)$$

$$C^{(0)} := \begin{pmatrix} e^{i(\chi_1 + \chi_2)} \cos \beta & -i e^{i\chi_1} t_+^x t_+^y e^{i\chi_4} \sin \beta \\ e^{i\chi_3} t_-^x t_-^y e^{i\chi_2} \sin \beta & i e^{i(\chi_3 + \chi_4)} \cos \beta \end{pmatrix}, \quad (34)$$

$$D^{(0)} := \begin{pmatrix} e^{i(\chi_2 + \chi_1)} \cos \beta & e^{i\chi_2} t_+^x t_+^y e^{i\chi_3} \sin \beta \\ -i e^{i\chi_4} t_-^x t_-^y e^{i\chi_1} \sin \beta & i e^{i(\chi_4 + \chi_3)} \cos \beta \end{pmatrix}. \quad (35)$$

Observe that in the limit $\theta = 0$, the \mathbb{Z}_2 network model reduces to two decoupled U(1) network models where each time evolution is essentially the same as the one for the U(1) network model derived in [32].

In the vicinity of the Chalker-Coddington quantum critical point (24), we find the 4×4 block diagonal Hamiltonian

$$\mathcal{H}_{SS'}^{(0)} = \begin{pmatrix} D_+ & 0 \\ 0 & D_- \end{pmatrix} \quad (36)$$

where the 2×2 block are expressed in terms of linear combinations of the 2×2 unit matrix σ_0 and of the Pauli matrices σ_x , σ_y , and σ_z according to

$$D_+ = \sigma_z (-i\partial_x + A_x) - \sigma_x (-i\partial_y + A_y) - \sigma_y m + \sigma_0 A_0, \quad (37)$$

and

$$D_- = -\sigma_y (-i\partial_x - A_x) + \sigma_z (-i\partial_y - A_y) + \sigma_x m + \sigma_0 A_0. \quad (38)$$

Thus, each 2×2 block Hamiltonian is of the Dirac form whereby the linear combinations

$$A_0 := -(\chi_1 + \chi_2 + \chi_3 + \chi_4), \quad (A_x, A_y) := (-\chi_1 + \chi_3, \chi_2 - \chi_4), \quad (39)$$

enter as a scalar gauge potential and a vector gauge potential would do, respectively.

Any deviation of θ from $\theta = 0$ lifts the reducibility of (36). To leading order in θ and close to the Chalker-Coddington quantum critical point (24),

$$\begin{aligned}\mathcal{M}_{SS'} &= \left(\mathcal{M}_S^{(0)} + \theta \mathcal{M}_S^{(1)} + \dots \right) \left(\mathcal{M}_{S'}^{(0)} + \theta \mathcal{M}_{S'}^{(1)} + \dots \right) \\ &= \mathcal{M}_{SS'}^{(0)} + \theta \left(\mathcal{M}_S^{(1)} \mathcal{M}_{S'}^{(0)} + \mathcal{M}_S^{(0)} \mathcal{M}_{S'}^{(1)} \right) + \dots\end{aligned}\quad (40)$$

with

$$\mathcal{M}_S^{(1)} = \begin{pmatrix} A^{(1)} & 0 \\ 0 & B^{(1)} \end{pmatrix}, \quad A^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad B^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad (41)$$

$$\mathcal{M}_{S'}^{(1)} = \begin{pmatrix} C^{(1)} & 0 \\ 0 & D^{(1)} \end{pmatrix}, \quad C^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad D^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (42)$$

where we have set $m = \chi_n = 0$ and $t_{\pm}^{x,y} = 1$. We obtain

$$\mathcal{H}_{SS'} = \begin{pmatrix} D_+ & D_\theta \\ D_\theta^\dagger & D_- \end{pmatrix}, \quad D_\theta := \theta \begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix} \quad (43)$$

to this order.

Next, we perform a sequence of unitary transformation generated by

$$U = \begin{pmatrix} e^{i\pi\sigma_y/4} & 0 \\ 0 & e^{i\pi\sigma_z/4} \end{pmatrix} \begin{pmatrix} e^{-i\pi\sigma_x/4} & 0 \\ 0 & e^{-i\pi\sigma_x/4} \end{pmatrix} \begin{pmatrix} e^{-i\pi/8} & 0 \\ 0 & e^{i\pi/8} \end{pmatrix}, \quad (44)$$

yielding

$$\mathcal{H} := U^\dagger \mathcal{H}_{SS'} U = \begin{pmatrix} \mathcal{H}_+ & \alpha\sigma_0 \\ \alpha\sigma_0 & \mathcal{H}_- \end{pmatrix} \quad (45)$$

with $\alpha = \sqrt{2}\theta$ and

$$\mathcal{H}_\pm = \sigma_x (-i\partial_x \pm A_x) + \sigma_y (-i\partial_y \pm A_y) \pm \sigma_z m + \sigma_0 A_0. \quad (46)$$

The 2×2 matrices \mathcal{H}_+ and \mathcal{H}_- describe a Dirac fermion with mass $\pm m$ in the presence of random vector potential $\pm(A_x, A_y)$ and random scalar potential A_0 , each of which is an effective Hamiltonian for the plateau transition of integer quantum Hall effect [32, 33]. The \mathcal{H}_\pm sectors are coupled by the matrix element $\alpha\sigma_0$.

The 4×4 continuum Dirac Hamiltonian \mathcal{H} can be written in the form

$$\begin{aligned}\mathcal{H} &= (-i\partial_x \sigma_x - i\partial_y \sigma_y) \otimes \tau_0 + (A_x \sigma_x + A_y \sigma_y + m\sigma_z) \otimes \tau_z \\ &\quad + A_0 \sigma_0 \otimes \tau_0 + \alpha \sigma_0 \otimes \tau_x,\end{aligned}\quad (47)$$

where τ_0 is a unit 2×2 matrix and τ_x , τ_y , and τ_z are three Pauli matrices. The Hamiltonian (47) is invariant for each realization of disorder under the operation

$$T \mathcal{H}^* T^{-1} = \mathcal{H}, \quad T := i\sigma_y \otimes \tau_x, \quad (48)$$

that implements time-reversal for a spin-1/2 particle.

The Dirac Hamiltonian (47) is the main result of this subsection. It is an effective model for the Anderson localization of quantum spin Hall systems, which belongs to

the symplectic class in view of the symmetry property (48). The Anderson transition in the Dirac Hamiltonian (47) should possess the same universal critical properties as those found in our numerical simulations of the \mathbb{Z}_2 network model. In the presence of the “Rashba” coupling α , there should appear a metallic phase near $m = 0$ which is surrounded by two insulating phases. In the limit $\alpha \rightarrow 0$, the metallic phase should shrink into a critical point of the integer quantum Hall plateau transition.

The 4×4 continuum Dirac Hamiltonian \mathcal{H} should be contrasted with a 2×2 Hamiltonian of a Dirac particle in random scalar potential,

$$\mathcal{H}_2 = -i\partial_x\sigma_x - i\partial_y\sigma_y + V(x, y)\sigma_0, \quad (49)$$

which has the minimal dimensionality of the Clifford algebra in $(2 + 1)$ -dimensional space time and is invariant under time-reversal operation, $\sigma_y\mathcal{H}_2^*\sigma_y = \mathcal{H}_2$. The 2×2 Dirac Hamiltonian (49) is an effective Hamiltonian for massless Dirac fermions on the surface of a three-dimensional \mathbb{Z}_2 topological insulator. After averaging over the disorder potential V , the problem of Anderson localization of the surface Dirac fermions is reduced to a NLSM with a \mathbb{Z}_2 topological term [25, 26]. Interestingly, this \mathbb{Z}_2 topological term prevents the surface Dirac fermions from localizing [27, 28]. It is this absence of two-dimensional localization that defines a three-dimensional \mathbb{Z}_2 topological insulator [29]. In contrast, the doubling of the size of the Hamiltonian (47) implies that the NLSM describing the Anderson localization in the 4×4 Hamiltonian (45) does not come with a \mathbb{Z}_2 topological term, because two \mathbb{Z}_2 topological terms cancel each other. We can thus conclude that the critical properties of metal-insulator transitions in the \mathbb{Z}_2 network model are the same as those in the standard symplectic class, in agreement with results of our numerical simulations of the \mathbb{Z}_2 network model [19, 20].

Before closing this subsection, we briefly discuss the Dirac Hamiltonian (47) in the clean limit where $A_0 = A_x = A_y = 0$. Since the system in the absence of disorder is translationally invariant, momentum is a good quantum number. We thus consider the Hamiltonian in momentum space

$$\mathcal{H}(\mathbf{k}) = \begin{pmatrix} k_x\sigma_x + k_y\sigma_y + m\sigma_z & \alpha\sigma_0 \\ \alpha\sigma_0 & k_x\sigma_x + k_y\sigma_y - m\sigma_z \end{pmatrix}, \quad (50)$$

where the wave number $\mathbf{k} = (k_x, k_y)$. When $\alpha = 0$, the Hamiltonian (50) becomes a direct sum of 2×2 Dirac Hamiltonian with mass of opposite signs. This is essentially the same low-energy Hamiltonian as the one appearing in the quantum spin Hall effect in HgTe/(Hg,Cd)Te quantum wells [6].

3.3. \mathbb{Z}_2 topological number

We now discuss the topological property of the time-reversal invariant insulator which is obtained from the effective Hamiltonian (50) of the \mathbb{Z}_2 network model in the absence of disorder. The topological attribute of the band insulator is intimately tied to the invariance

$$\hat{\Theta}^{-1}\mathcal{H}(-\mathbf{k})\hat{\Theta} = \mathcal{H}(\mathbf{k}) \quad (51)$$

under the operation of time-reversal represented by

$$\hat{\Theta} := (i\sigma_y \otimes \tau_x)\mathcal{K} = -\hat{\Theta}^{-1}, \quad (52)$$

where \mathcal{K} implements complex conjugation. We are going to show that this topological attribute takes values in \mathbb{Z}_2 , i.e., the \mathbb{Z}_2 index introduced by Kane and Mele [4].

We begin with general considerations on a translation-invariant single-particle fermionic Hamiltonian which has single-particle eigenstates labeled by the wave vector \mathbf{k} taking values in a compact manifold. This compact manifold can be the first Brillouin zone with the topology of a torus if the Hamiltonian is defined on a lattice and periodic boundary conditions are imposed, or it can be the stereographic projection between the momentum plane \mathbb{R}^2 and the surface of a three-dimensional sphere if the Hamiltonian is defined in the continuum. We assume that (i) the antiunitary operation $\hat{\Theta} = -\hat{\Theta}^{-1} = -\hat{\Theta}^\dagger$ that implements time-reversal leaves the Hamiltonian invariant, (ii) there exists a spectral gap at the Fermi energy, and (iii) there are two distinct occupied bands with the single-particle orthonormal eigenstates $|u_{\hat{a}}(\mathbf{k})\rangle$ and energies $E_{\hat{a}}(\mathbf{k})$ labeled by the index $\hat{a} = 1, 2$ below the Fermi energy. All three assumptions are met by the 4×4 Dirac Hamiltonian (47), provided that the mass m is nonvanishing.

Because of assumptions (i) and (ii) the 2×2 unitary sewing matrix with the matrix elements $w_{\hat{a}\hat{b}}(\mathbf{k})$ defined by

$$w_{\hat{a}\hat{b}}(\mathbf{k}) := \langle u_{\hat{a}}(-\mathbf{k}) | \left(\hat{\Theta} | u_{\hat{b}}(\mathbf{k}) \rangle \right) \equiv \langle u_{\hat{a}}(-\mathbf{k}) | \Theta u_{\hat{b}}(\mathbf{k}) \rangle, \quad \hat{a}, \hat{b} = 1, 2, \quad (53)$$

i.e., the overlaps between the occupied single-particle energy eigenstates with momentum $-\mathbf{k}$ and the time reversed images to the occupied single-particle energy eigenstates with momentum \mathbf{k} , plays an important role [17]. The matrix elements (53) obey

$$\begin{aligned} w_{\hat{a}\hat{b}}(\mathbf{k}) &\equiv \langle u_{\hat{a}}(-\mathbf{k}) | \left(\hat{\Theta} | u_{\hat{b}}(\mathbf{k}) \rangle \right) \\ &= \langle u_{\hat{b}}(\mathbf{k}) | \left(\hat{\Theta}^\dagger | u_{\hat{a}}(-\mathbf{k}) \rangle \right) \\ &= -\langle u_{\hat{b}}(\mathbf{k}) | \left(\hat{\Theta} | u_{\hat{a}}(-\mathbf{k}) \rangle \right) \\ &\equiv -w_{\hat{b}\hat{a}}(-\mathbf{k}), \end{aligned} \quad \hat{a}, \hat{b} = 1, 2. \quad (54)$$

We used the fact that $\hat{\Theta}$ is antilinear to reach the second equality and that it is antiunitary with $\hat{\Theta}^2 = -1$ to reach the third equality. Hence, the 2×2 unitary sewing matrix $w(\mathbf{k})$ with the matrix elements (53) can be parametrized as

$$w(\mathbf{k}) = \begin{pmatrix} w_{11}(\mathbf{k}) & w_{12}(\mathbf{k}) \\ -w_{12}(-\mathbf{k}) & w_{22}(\mathbf{k}) \end{pmatrix} = -w^T(-\mathbf{k}) \quad (55)$$

with the three complex-valued functions

$$w_{11}(\mathbf{k}) = -w_{11}(-\mathbf{k}), \quad w_{22}(\mathbf{k}) = -w_{22}(-\mathbf{k}), \quad w_{12}(\mathbf{k}). \quad (56)$$

We observe that $w(\mathbf{k})$ reduces to

$$w(\mathbf{k}) = e^{if(\mathbf{k})} \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix} \quad (57)$$

for some real-valued $f(\mathbf{k})$ at any time-reversal invariant wave vector $\mathbf{k} \sim -\mathbf{k}$ (time-reversal invariant wave vectors are half a reciprocal vector for a lattice model, and 0 or ∞ for a model in the continuum).

As we shall shortly see, the sewing matrix (53) imposes constraints on the U(2) Berry connection

$$\mathcal{A}_{\hat{a}\hat{b}}(\mathbf{k}) := \langle u_{\hat{a}}(\mathbf{k}) | du_{\hat{b}}(\mathbf{k}) \rangle \equiv \left\langle u_{\hat{a}}(\mathbf{k}) \left| \frac{\partial}{\partial k_{\mu}} u_{\hat{b}}(\mathbf{k}) \right. \right\rangle dk_{\mu} \equiv A_{\hat{a}\hat{b}}^{\mu}(\mathbf{k}) dk_{\mu}, \quad (58)$$

where the summation convention over the repeated index μ is understood (we do not make distinction between superscript and subscript). Here, at every point \mathbf{k} in momentum space, we have introduced the U(2) antihermitian gauge field $A_{\mu}(\mathbf{k})$ with the space index $\mu = 1, 2$ and the matrix elements

$$A_{\hat{a}\hat{b}}^{\mu}(\mathbf{k}) = - \left(A_{\hat{b}\hat{a}}^{\mu}(\mathbf{k}) \right)^* \quad (59)$$

labeled with the U(2) internal indices $\hat{a}, \hat{b} = 1, 2$, by performing an infinitesimal parametric change in the Hamiltonian. We decompose the U(2) gauge field (58) into the U(1) and the SU(2) contributions

$$A_{\mu}(\mathbf{k}) \equiv a_{\mu}^0(\mathbf{k}) \frac{\rho_0}{2i} + \boldsymbol{a}_{\mu}(\mathbf{k}) \cdot \frac{\boldsymbol{\rho}}{2i}, \quad (60)$$

where ρ_0 is a 2×2 unit matrix and $\boldsymbol{\rho}$ is a 3 vector made of the Pauli matrices ρ_x, ρ_y , and ρ_z . Accordingly,

$$\mathcal{A}^{\text{U}(2)}(\mathbf{k}) = \mathcal{A}^{\text{U}(1)}(\mathbf{k}) + \mathcal{A}^{\text{SU}(2)}(\mathbf{k}). \quad (61)$$

Combining the identity $\hat{\Theta}^2 = -1$ with the (partial) resolution of the identity $\sum_{\hat{a}=1,2} |u_{\hat{a}}(\mathbf{k})\rangle \langle u_{\hat{a}}(\mathbf{k})|$ for the occupied energy eigenstates with momentum \mathbf{k} yields

$$\sum_{\hat{a}=1,2} \hat{\Theta} |u_{\hat{a}}(\mathbf{k})\rangle \langle u_{\hat{a}}(\mathbf{k})| \hat{\Theta} = -1, \quad (62)$$

where the proper restriction to the occupied energy eigenstates is understood for the unit operator on the right-hand side. Using this identity, we deduce the gauge transformation

$$\begin{aligned} A_{\mu}(-\mathbf{k}) &= - \left(\left\langle u_{\hat{a}}(-\mathbf{k}) \left| \frac{\partial}{\partial k_{\mu}} u_{\hat{b}}(-\mathbf{k}) \right. \right\rangle \right)_{\hat{a}, \hat{b}=1,2} \\ &= -w(\mathbf{k}) A_{\mu}^*(\mathbf{k}) w^{\dagger}(\mathbf{k}) - w(\mathbf{k}) \partial_{\mu} w^{\dagger}(\mathbf{k}) \\ &= +w(\mathbf{k}) A_{\mu}^{\text{T}}(\mathbf{k}) w^{\dagger}(\mathbf{k}) - w(\mathbf{k}) \partial_{\mu} w^{\dagger}(\mathbf{k}) \end{aligned} \quad (63)$$

that relates the U(2) connections at $\pm\mathbf{k}$. For the U(1) and SU(2) parts of the connection,

$$a_{\mu}^0(-\mathbf{k}) = a_{\mu}^0(\mathbf{k}) - 2\partial_{\mu}\zeta(\mathbf{k}), \quad (64)$$

$$\boldsymbol{a}_{\mu}(-\mathbf{k}) \cdot \boldsymbol{\rho} = \boldsymbol{a}_{\mu}(\mathbf{k}) \cdot \tilde{w}(\mathbf{k}) \boldsymbol{\rho}^{\text{T}} \tilde{w}^{\dagger}(\mathbf{k}) - 2i \tilde{w}(\mathbf{k}) \partial_{\mu} \tilde{w}^{\dagger}(\mathbf{k}), \quad (65)$$

where we have decomposed $w(\mathbf{k})$ into the U(1) ($e^{i\zeta}$) and SU(2) (\tilde{w}) parts according to

$$w(\mathbf{k}) = e^{i\zeta(\mathbf{k})} \tilde{w}(\mathbf{k}), \quad (66)$$

(note that this decomposition has a global sign ambiguity, which, however, will not affect the following discussions).

Equipped with these gauge fields, we introduce the U(2) Wilson loop

$$\begin{aligned} W_{\text{U}(2)}[\mathcal{C}] &:= \frac{1}{2} \text{tr} \mathcal{P} \exp \left(\oint_{\mathcal{C}} \mathcal{A}^{\text{U}(2)}(\mathbf{k}) \right) \\ &= W_{\text{U}(1)}[\mathcal{C}] \times W_{\text{SU}(2)}[\mathcal{C}], \end{aligned} \quad (67)$$

where the U(1) Wilson loop is given by

$$W_{\text{U}(1)}[\mathcal{C}] := \exp \left(\oint_{\mathcal{C}} \mathcal{A}^{\text{U}(1)}(\mathbf{k}) \right), \quad (68)$$

while the SU(2) Wilson loop is given by

$$W_{\text{SU}(2)}[\mathcal{C}] := \frac{1}{2} \text{tr} \mathcal{P} \exp \left(\oint_{\mathcal{C}} \mathcal{A}^{\text{SU}(2)}(\mathbf{k}) \right). \quad (69)$$

The symbol \mathcal{P} in the definition of the U(2) Wilson loop represents path ordering, while \mathcal{C} is any closed loop in the compact momentum space.

By construction, the U(2) Wilson loop (67) is invariant under the transformation

$$A^\mu(\mathbf{k}) \rightarrow U^\dagger(\mathbf{k}) A^\mu(\mathbf{k}) U(\mathbf{k}) + U^\dagger(\mathbf{k}) \partial^\mu U(\mathbf{k}) \quad (70)$$

induced by the local (in momentum space) U(2) transformation

$$|u_{\hat{a}}(\mathbf{k})\rangle \rightarrow |u_{\hat{b}}(\mathbf{k})\rangle U_{\hat{b}\hat{a}}(\mathbf{k}) \quad (71)$$

on the single-particle energy eigenstates. Similarly, the SU(2) and U(1) Wilson loops are invariant under any local SU(2) and U(1) gauge transformation of the Bloch wave functions, respectively.

When \mathcal{C} is invariant as a set under

$$\mathbf{k} \rightarrow -\mathbf{k}, \quad (72)$$

the SU(2) Wilson loop $W_{\text{SU}(2)}[\mathcal{C}]$ is quantized to the two values

$$W_{\text{SU}(2)}[\mathcal{C}] = \pm 1 \quad (73)$$

because of time-reversal symmetry. Furthermore, the identity

$$W_{\text{SU}(2)}[\mathcal{C}] = \prod_{\substack{\mathbf{K} \sim -\mathbf{K} \\ \mathbf{K} \in \mathcal{C}}} \text{Pf}(\tilde{w}(\mathbf{K})), \quad (74)$$

which we will prove below, follows. Here, the symbol Pf denotes the Pfaffian of an antisymmetric matrix, and only the subset of momenta $\mathbf{K} \in \mathcal{C}$ that are unchanged under $\mathbf{K} \rightarrow -\mathbf{K}$ contribute to the SU(2) Wilson loop. According to (54), the sewing matrix at a time-reversal symmetric wave vector is an antisymmetric 2×2 matrix. Consequently, the SU(2) part of the sewing matrix at a time-reversal symmetric wave vector is a real-valued antisymmetric 2×2 matrix (i.e., it is proportional to $i\rho_y$ up to a sign). Hence, its Pfaffian is a well-defined and nonvanishing real-valued number.

Before undertaking the proof of (74), more insights on this identity can be obtained if we specialize to the case when the Hamiltonian is invariant under any U(1) subgroup

of SU(2), e.g., the z -component of spin σ_z . In this case we can choose the basis states which diagonalize σ_z ; $\sigma_z|u_1(\mathbf{k})\rangle = +|u_1(\mathbf{k})\rangle$, $\sigma_z|u_2(\mathbf{k})\rangle = -|u_2(\mathbf{k})\rangle$. Since the time-reversal operation changes the sign of σ_z , the sewing matrix takes the form

$$w(\mathbf{k}) = \begin{pmatrix} 0 & e^{-i\chi(\mathbf{k})} \\ -e^{-i\chi(-\mathbf{k})} & 0 \end{pmatrix}, \quad (75)$$

which, in combination with (64) and (65) implies the transformation laws

$$a_\mu^0(-\mathbf{k}) = +a_\mu^0(\mathbf{k}) + \partial_\mu [\chi(\mathbf{k}) + \chi(-\mathbf{k})], \quad (76)$$

$$a_\mu^z(-\mathbf{k}) = -a_\mu^z(\mathbf{k}) + \partial_\mu [\chi(\mathbf{k}) - \chi(-\mathbf{k})]. \quad (77)$$

We conclude that when both the z component of the electron spin and the electron number are conserved, we can set

$$a_\mu^x(\mathbf{k}) = a_\mu^y(\mathbf{k}) = 0, \quad A_\mu^{\text{U}(2)}(\mathbf{k}) = a_\mu^0(\mathbf{k}) \frac{\sigma_0}{2i} + a_\mu^z(\mathbf{k}) \frac{\sigma_z}{2i}, \quad (78)$$

and use the transformation law

$$A_{\nu,11}^{\text{U}(2)}(-\mathbf{k}) = \frac{1}{2i} [a_\nu^0(-\mathbf{k}) + a_\nu^z(-\mathbf{k})] = A_{\nu,22}^{\text{U}(2)}(\mathbf{k}) - i\partial_\nu \chi(\mathbf{k}). \quad (79)$$

With conservation of the z component of the electron spin in addition to that of the electron charge, the SU(2) Wilson loop becomes

$$W_{\text{SU}(2)}[\mathcal{C}] = \frac{1}{2} \text{tr} \mathcal{P} \exp \left(\oint_{\mathcal{C}} \mathcal{A}^{\text{SU}(2)}(\mathbf{k}) \right) \quad (80)$$

$$\begin{aligned} &= \frac{1}{2} \text{tr} \exp \left(\oint_{\mathcal{C}} a_\mu^z(\mathbf{k}) \frac{\sigma_z}{2i} dk^\mu \right) \\ &= \cos \left(\frac{1}{2} \oint_{\mathcal{C}} a_\mu^z(\mathbf{k}) dk^\mu \right). \end{aligned} \quad (81)$$

We have used the fact that σ_z is traceless to reach the last line. This line integral can be written as the surface integral

$$\oint_{\mathcal{C}} a_\mu^z(\mathbf{k}) dk^\mu = \int_{\mathcal{D}} d^2k \varepsilon^{\mu\nu} \partial_\mu a_\nu^z(\mathbf{k}) \quad (82)$$

by Stokes' theorem. Here, \mathcal{D} is the region defined by $\partial\mathcal{D} = \mathcal{C}$, and covers a half of the total Brillouin zone (BZ) because of the condition (72). In turn, this surface integral is equal to the Chern number for up-spin fermions,

$$\begin{aligned} \text{Ch}_\uparrow &:= \int_{\text{BZ}} \frac{d^2k}{2\pi i} \varepsilon^{\mu\nu} \partial_\mu A_{\nu,11}^{\text{U}(2)}(\mathbf{k}) \\ &\equiv \int_{\text{BZ}} \frac{d^2k}{2\pi i} F_{11}^{\text{U}(2)}(\mathbf{k}) \\ &= \int_{\mathcal{D}} \frac{d^2k}{2\pi i} [F_{11}^{\text{U}(2)}(\mathbf{k}) + F_{11}^{\text{U}(2)}(-\mathbf{k})] \end{aligned} \quad (83)$$

$$\begin{aligned}
&= \int_{\mathcal{D}} \frac{d^2 k}{2\pi i} \varepsilon^{\mu\nu} \partial_\mu \left[A_{\nu,11}^{U(2)}(\mathbf{k}) - A_{\nu,22}^{U(2)}(\mathbf{k}) \right] \\
&= -i \int_{\mathcal{D}} \frac{d^2 k}{2\pi i} \varepsilon^{\mu\nu} \partial_\mu a_\nu^z(\mathbf{k}),
\end{aligned} \tag{84}$$

where we have used the transformation law (79) to deduce that

$$F_{11}^{U(2)}(-\mathbf{k}) = -F_{22}^{U(2)}(\mathbf{k}) \tag{85}$$

to reach the fourth equality.

To summarize, when the z component of the spin is conserved, the quantized $SU(2)$ Wilson loop can then be written as the parity of the spin Chern number (the Chern number for up-spin fermions, which is equal to minus the Chern number for down-spin fermions) [3, 4, 5],

$$W_{SU(2)}[\mathcal{C}] = (-1)^{\text{Ch}\uparrow}. \tag{86}$$

Next, we apply the master formula (74) to the 4×4 Dirac Hamiltonian (50). To this end, we first replace the mass m by the k -dependent mass,

$$m_k = m - C\mathbf{k}^2, \quad C > 0, \tag{87}$$

and parametrize the wave number \mathbf{k} as

$$k_x + ik_y = k e^{i\varphi}, \quad -\infty < k < \infty, \quad 0 \leq \varphi < \pi. \tag{88}$$

Without loss of generality, we may assume $\alpha > 0$. The mass m_k is introduced so that the $SU(2)$ part of the sewing matrix is single-valued in the limit $|k| \rightarrow \infty$.

We then perform another series of unitary transformation with

$$\tilde{U} = \begin{pmatrix} \sigma_0 & 0 \\ 0 & i\sigma_z \end{pmatrix} \begin{pmatrix} \frac{\sigma_0}{\sqrt{2}} & \frac{\sigma_0}{\sqrt{2}} \\ -\frac{\sigma_0}{\sqrt{2}} & \frac{\sigma_0}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} e^{i\pi\sigma_z/4} & 0 \\ 0 & e^{-i\pi\sigma_z/4} \end{pmatrix}, \tag{89}$$

to rewrite the Hamiltonian (50) in the form

$$\begin{aligned}
\tilde{\mathcal{H}}(\mathbf{k}) &:= \tilde{U}^\dagger \mathcal{H}(\mathbf{k}) \tilde{U} \\
&= \begin{pmatrix} 0 & k_x \sigma_x + k_y \sigma_y + (\alpha - im_k) \sigma_0 \\ k_x \sigma_x + k_y \sigma_y + (\alpha + im_k) \sigma_0 & 0 \end{pmatrix}.
\end{aligned} \tag{90}$$

The four eigenvalues of the Hamiltonian (90) are given by $E(\mathbf{k}) = \pm\lambda_k^+, \pm\lambda_k^-$, where

$$\lambda_k^\pm = \sqrt{(k \pm \alpha)^2 + m_k^2}. \tag{91}$$

The occupied eigenstate with the energy $E_1(\mathbf{k}) = -\lambda_k^-$ reads

$$|u_1(\varphi, k)\rangle = \frac{1}{2\lambda_k^-} \begin{pmatrix} -\lambda_k^- \\ \lambda_k^- e^{-i\varphi} \\ -k + \alpha + im_k \\ (k - \alpha - im_k) e^{-i\varphi} \end{pmatrix}, \tag{92}$$

Table 1. θ_k^\pm at the time-reversal invariant momenta $k = 0$ and $k = \pm\infty$, when $m - C\alpha^2 < 0$ (a) and $m - C\alpha^2 > 0$ (b). It is assumed that $0 \leq \arctan(|m|/\alpha) \leq \pi/2$.

(a) $m - C\alpha^2 < 0$				
k	$-\infty$	0	$+\infty$	
θ_k^+	$\pi/2$	$\arctan(-m/\alpha)$	$\pi/2$	
θ_k^-	$-\pi/2$	$\arctan(-m/\alpha) - \pi$	$-\pi/2$	
$e^{i(\theta_k^+ - \theta_k^-)/2}$	i	i	i	
(b) $m - C\alpha^2 > 0$				
k	$-\infty$	0	$+\infty$	
θ_k^+	$-3\pi/2$	$-\arctan(m/\alpha)$	$\pi/2$	
θ_k^-	$3\pi/2$	$\pi - \arctan(m/\alpha)$	$-\pi/2$	
$e^{i(\theta_k^+ - \theta_k^-)/2}$	i	$-i$	i	

and the occupied eigenstate with the energy $E_2(\mathbf{k}) = -\lambda_k^+$ is

$$|u_2(\varphi, k)\rangle = \frac{1}{2\lambda_k^+} \begin{pmatrix} -\lambda_k^+ \\ -\lambda_k^+ e^{-i\varphi} \\ k + \alpha + im_k \\ (k + \alpha + im_k)e^{-i\varphi} \end{pmatrix}. \quad (93)$$

Notice that $|u_2(\varphi, k)\rangle = |u_1(\varphi + \pi, -k)\rangle$.

The 2×2 sewing matrix $w(\mathbf{k})$ is obtained from the eigenstates (92)–(93) as

$$\begin{aligned} w(\varphi, k) &:= \left(\langle u_{\hat{a}}(\varphi, -k) | \hat{\Theta} | u_{\hat{b}}(\varphi, k) \rangle \right)_{\hat{a}, \hat{b}=1,2} \\ &= -e^{i\varphi} \begin{pmatrix} 0 & \frac{1}{\lambda_k^+}(k + \alpha - im_k) \\ \frac{1}{\lambda_k^-}(k - \alpha + im_k) & 0 \end{pmatrix}, \end{aligned} \quad (94)$$

which is decomposed into the U(1) part,

$$i \exp \left(i\varphi + i(\theta_k^+ + \theta_k^-)/2 \right), \quad (95)$$

and the SU(2) part,

$$\tilde{w}(k) = \begin{pmatrix} 0 & ie^{i(\theta_k^+ - \theta_k^-)/2} \\ ie^{-i(\theta_k^+ - \theta_k^-)/2} & 0 \end{pmatrix}, \quad (96)$$

of the sewing matrix. Here, we have defined θ_k^\pm through the relation

$$e^{i\theta_k^\pm} = \frac{1}{\lambda_k^\pm} [k \pm (\alpha - im_k)]. \quad (97)$$

For the SU(2) sewing matrix (96), there are two momenta which are invariant under inversion $\mathbf{k} \rightarrow -\mathbf{k}$, namely the south $\mathbf{K} = 0$ and north $\mathbf{K} = \infty$ poles of the stereographic sphere. The values of θ_k^\pm at these time-reversal momenta are listed in table 1. The



Figure 3. (a) Quantum spin Hall droplet immersed in the reference vacuum [in real space $(x, y) \in \mathbb{R}^2$]. (b) The \mathbb{Z}_2 network model or its tight-binding equivalent when $x < 0$ is separated from the reference vacuum at $x > 0$ by the vertical boundary $x = 0$ [in real space $(x, y) \in \mathbb{R}^2$].

Pfaffian of the sewing matrix at the south and north poles of the stereographic sphere are

$$\text{Pf } \tilde{w}(0) = -\text{sgn}(m - C\alpha^2) \text{Pf } (-i\rho_y), \quad (98)$$

$$\text{Pf } \tilde{w}(\infty) = \text{Pf } (-i\rho_y), \quad (99)$$

respectively. Hence,

$$W_{\text{SU}(2)}[\mathcal{C}] = -\text{sgn}(m) \quad (100)$$

for any time-reversal invariant path \mathcal{C} passing through the south and north poles, where we have suppressed $C\alpha^2$ by taking the limit $C\alpha^2/|m| \rightarrow 0$.

The value (100) taken by the $\text{SU}(2)$ Wilson loop thus appears to be ambiguous since it depends on the sign of the mass m . This ambiguity is a mere reflection of the fact that, as noted in [20], the topological nature of the \mathbb{Z}_2 network model is itself defined relative to that of some reference vacuum. Indeed, for any given choice of the parameters (X, θ) from figure 2 that defines uniquely the bulk properties of the insulating phase in the \mathbb{Z}_2 network model, the choice of boundary conditions determines if a single helical Kramers' doublet edge state is or is not present at the boundary of the \mathbb{Z}_2 network model. In view of this, it is useful to reinterpret the \mathbb{Z}_2 network model with a boundary as realizing a quantum spin Hall droplet immersed in a reference vacuum as is depicted in figure 3(a). If so, choosing the boundary condition is equivalent to fixing the topological attribute of the reference vacuum *relative to* that of the \mathbb{Z}_2 network model, for the reference vacuum in which the quantum spin Hall droplet is immersed also has either a trivial or non-trivial \mathbb{Z}_2 quantum topology. A single helical Kramers' doublet propagating unhindered along the boundary between the quantum spin Hall droplet and the reference vacuum appears if and only if the \mathbb{Z}_2 topological quantum numbers in the droplet and in the reference vacuum differ.

In the low-energy continuum limit (50), a boundary in real space can be introduced by breaking translation invariance along the vertical line $x = 0$ in the real space

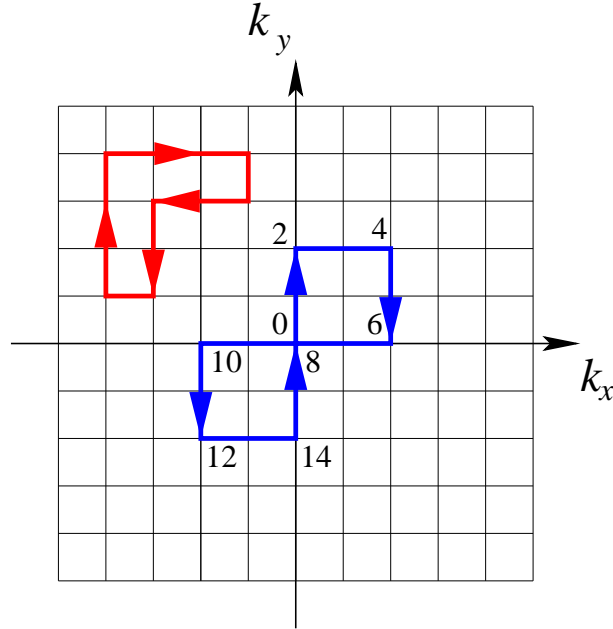


Figure 4. Momentum space $(k_x, k_y) \in \mathbb{R}^2$ is discretized with the help of a rectangular grid on which two paths are depicted. The red path that is restricted to the upper left quadrant is not invariant as a set under the inversion $(k_x, k_y) \rightarrow -(k_x, k_y)$. The blue path with its center of mass at the origin is. This path is assembled out of 16 links: $(i_0, i_1) = -(i_{15}, i_0)$, $(i_1, i_2) = -(i_{14}, i_1)$, $(i_2, i_3) = -(i_{13}, i_2)$, $(i_3, i_4) = -(i_{12}, i_3)$, $(i_4, i_5) = -(i_{11}, i_4)$, $(i_5, i_6) = -(i_{10}, i_5)$, $(i_6, i_7) = -(i_9, i_6)$, $(i_7, i_8) = -(i_8, i_7)$. Sites $i_0 = i_8 = i_{16}$ along the path are the only ones invariant under $(k_x, k_y) \rightarrow -(k_x, k_y)$.

$(x, y) \in \mathbb{R}^2$ through the profile [see figure 3(b)]

$$m(x, y) = m(x) = \begin{cases} -m, & \text{if } x \rightarrow -\infty, \\ +m, & \text{if } x \rightarrow +\infty, \end{cases} \quad (101)$$

for the mass.

We close Sec. 3 with a justification of the master formula (74). To this end, we regularize the continuum gauge theory by discretizing momentum space (figure 4). We use the momentum coordinate $i \in \mathbb{Z}^2$ on a rectangular grid with the two lattice spacings $\Delta k^\mu > 0$. To each link from the site i to the nearest-neighbor site $i + \mu$ of the grid, we assign the SU(2) unitary matrix

$$U_{i, i+\mu} \equiv e^{A_{i, i+\mu} \Delta k^\mu}, \quad (102)$$

which is obtained by discarding U(1) part of the U(2) Berry connection. Consistency demands that

$$U_{i+\mu, i} = U_{i, i+\mu}^\dagger \iff A_{i+\mu, i} = A_{i, i+\mu}^\dagger. \quad (103)$$

We define the SU(2) Wilson loop to be

$$W_{\text{SU}(2)}(i_0, \dots, i_{N-1}) := \frac{1}{2} \text{tr} \left(U_{i_0, i_1} U_{i_1, i_2} \cdots U_{i_{N-1}, i_0} \right) \quad (104)$$

where i_n and i_{n+1} are nearest neighbors, i.e., their difference $i_{n+1} - i_n = \eta_n$ is a unit vector η_n . The Wilson loop is invariant under any local gauge transformation by which

$$U_{i,i+\mu} \rightarrow V_i^\dagger U_{i,i+\mu} V_{i+\mu} \quad (105)$$

where the V_i 's are $U(2)$ matrices. Observe that the cyclicity of the trace allows us to write

$$W_{\text{SU}(2)}(i_0, \dots, i_{N-1}) = \frac{1}{2} \text{tr} \left(U_{i_{\frac{N}{2}}, i_{\frac{N}{2}+1}} \cdots U_{i_{N-1}, i_0} U_{i_0, i_1} U_{i_1, i_2} \cdots U_{i_{\frac{N}{2}-1}, i_{\frac{N}{2}}} \right). \quad (106)$$

To make contact with the master formula (74), we assume that the closed path with vertices i_ℓ parametrized by the index $\ell = 0, 1, \dots, N-1$ obeys the condition that

$$\begin{aligned} & \vdots \\ & i_{N-n} \text{ is the wave vector } -\sum_{m=1}^n \eta_m \Delta k^{\mu_m}, \\ & \vdots \\ & i_{N-1} \text{ is the wave vector } -\eta_1 \Delta k^{\mu_1}, \\ & i_0 \text{ is the wave vector } 0, \\ & i_1 \text{ is the wave vector } +\eta_1 \Delta k^{\mu_1}, \\ & \vdots \\ & i_n \text{ is the wave vector } +\sum_{m=1}^n \eta_m \Delta k^{\mu_m}, \\ & \vdots \end{aligned} \quad (107)$$

with $\eta_m = \pm 1$ for $m = 1, \dots, N/2$ in order to mimic after discretization the condition that the closed path entering the Wilson loop is invariant as a set under the inversion (72).

On the discretized momentum lattice the sewing matrix (53) is defined by

$$(w_i)_{\hat{a}\hat{b}} := \langle u_{\hat{a}}(-i) | \hat{\Theta} | u_{\hat{b}}(i) \rangle, \quad (108)$$

which obeys the condition

$$w_{-i} = -w_i^T, \quad (109)$$

i.e., the counterpart to the relation (54). This implies that w_{i_0} and $w_{i_{N/2}}$ are antisymmetric unitary 2×2 matrices. Furthermore, the sewing matrix w_i must also obey the counterpart to (63), namely

$$U_{-j, -i} = w_j U_{i,j}^T w_i^\dagger. \quad (110)$$

It now follows from (107) and (110) that

$$\begin{aligned} & U_{i_{N-1}, i_0} = w_{i_1} U_{i_0, i_1}^T w_{i_0}^\dagger, \\ & \vdots \\ & U_{i_{N-1-n}, i_{N-n}} = w_{i_{n+1}} U_{i_n, i_{n+1}}^T w_{i_n}^\dagger, \\ & \vdots \\ & U_{i_{\frac{N}{2}}, i_{\frac{N}{2}+1}} = w_{i_{\frac{N}{2}}} U_{i_{\frac{N}{2}-1}, i_{\frac{N}{2}}}^T w_{i_{\frac{N}{2}-1}}^\dagger. \end{aligned} \quad (111)$$

In particular, we observe that

$$U_{i_{N-1},i_0} U_{i_0,i_1} = w_{i_1} U_{i_0,i_1}^T w_{i_0}^\dagger U_{i_0,i_1} = w_{i_1} w_{i_0}^\dagger U_{i_0,i_1}^T U_{i_0,i_1} = w_{i_1} w_{i_0}^\dagger, \quad (112)$$

since w_{i_0} is a 2×2 antisymmetric unitary matrix, i.e., w_{i_0} is the second Pauli matrix up to a phase factor, while

$$(\boldsymbol{\rho} \cdot \mathbf{n})^T \rho_2 = -\rho_2 (\boldsymbol{\rho} \cdot \mathbf{n}) \quad (113)$$

holds for any three-vector \mathbf{n} contracted with the three-vector $\boldsymbol{\rho}$ made of the three Pauli matrices. By repeating the same exercise a second time,

$$\begin{aligned} U_{i_{N-2},i_{N-1}} \left(U_{i_{N-1},i_0} U_{i_0,i_1} \right) U_{i_1,i_2} &= w_{i_2} U_{i_1,i_2}^T w_{i_1}^\dagger \left(w_{i_1} w_{i_0}^\dagger \right) U_{i_1,i_2} \\ &= w_{i_2} w_{i_0}^\dagger, \end{aligned} \quad (114)$$

one convinces oneself that the dependences on the gauge fields A_{i_0,i_1} and A_{i_{N-1},i_0} , A_{i_1,i_2} and $A_{i_{N-2},i_{N-1}}$, and so on until A_{i_{n-1},i_n} and $A_{i_{N-n},i_{N-n+1}}$ at the level n of this iteration cancel pairwise due to the conditions (107)–(110) implementing time-reversal invariance. This iteration stops when $n = N/2$, in which case the $SU(2)$ Wilson loop is indeed solely controlled by the sewing matrix at the time-reversal invariant momenta corresponding to $\ell = 0$ and $\ell = N/2$,

$$W_{SU(2)}(i_0, \dots, i_{N-1}) = \frac{1}{2} \text{tr} \left(w_{i_{N/2}} w_{i_0}^\dagger \right). \quad (115)$$

Since i_0 and $i_{N/2}$ are invariant under momentum inversion or, equivalently, time-reversal invariant,

$$w_{i_{N/2}} = e^{i\alpha_{N/2}} i\rho_2, \quad w_{i_0} = e^{i\alpha_0} i\rho_2 \quad (116)$$

with $\alpha_{N/2}, \alpha_0 = 0, \pi$. Here, the \mathbb{Z}_2 phases $e^{i\alpha_{N/2}}$ and $e^{i\alpha_0}$ are none other than the Pfaffians

$$e^{i\alpha_{N/2}} = \text{Pf} \left(w_{i_{N/2}} \right), \quad e^{i\alpha_0} = \text{Pf} \left(w_{i_0} \right), \quad (117)$$

respectively. Hence,

$$\begin{aligned} W_{SU(2)}(i_0, \dots, i_{N-1}) &= \frac{1}{2} \text{tr} \left[\text{Pf} \left(w_{i_{N/2}} \right) i\rho_2 \times \text{Pf} \left(w_{i_0}^\dagger \right) (-i\rho_2) \right] \\ &= \text{Pf} \left(w_{i_{N/2}} \right) \text{Pf} \left(w_{i_0} \right) \end{aligned} \quad (118)$$

is a special case of (74). (Recall that w_{i_0} and $w_{i_{N/2}}$ are real-valued.)

4. Numerical study of boundary multifractality in the \mathbb{Z}_2 network model

In [20], we have shown that (i) multifractal scaling holds near the boundary of the \mathbb{Z}_2 network model at the transition between the metallic phase and the \mathbb{Z}_2 topological insulating phase shown in figure 2, (ii) it is different from that in the ordinary symplectic class, while (iii) bulk properties, such as the critical exponents for the divergence of the localization length and multifractal scaling in the bulk, are the same as those in the conventional two-dimensional symplectic universality class of Anderson localization. This implies that the boundary critical properties are affected by the presence of the

helical edge states in the topological insulating phase adjacent to the critical point. In this work, we improve the precision for the estimate of the boundary multifractal critical exponents. We also compute numerically additional critical exponents that encode corner (zero-dimensional) multifractality at the metal-to- \mathbb{Z}_2 -topological-insulator transition. We thereby support the claim that conformal invariance is present at the metal-to- \mathbb{Z}_2 -topological-insulator transition by verifying that conformal relations between critical exponents at these boundaries hold.

4.1. Boundary and corner multifractality

To characterize multifractal scaling at the metal-insulator transition in the \mathbb{Z}_2 network model, we start from the time-evolution of the plane waves along the links of the network with the scattering matrices defined in (1)-(7) at the nodes S and S' . To minimize finite size effects, the parameter θ in (1)-(7) is chosen to be a random variable as explained in Sec. 2. We focus on the metal-insulator transition at $X = X_l = 0.971$ as shown in figure 2(b).

When we impose reflecting boundary conditions, a node on the boundary reduces to a unit 2×2 matrix. When the horizontal reflecting boundaries are located at nodes of type S' , as shown in figure 5(a), there exists a single helical edge states for $X > X_l$. The insulating phase $X > X_l$ is thus topologically nontrivial.

For each realization of the disorder, we numerically diagonalize the one-step time-evolution operator of the \mathbb{Z}_2 network model and retain the normalized wave function $\psi_\sigma(x, y)$, after coarse graining over the 4 edges of the plaquette located at (x, y) , whose eigenvalue is the closest to 1. The wave function at criticality is observed to display the power-law dependence on the linear dimension L of the system,

$$\sum_{\sigma=\uparrow,\downarrow} |\psi_\sigma(x, y)|^{2q} \propto L^{-\Delta_q^{(\zeta,\nu)} - dq}. \quad (119)$$

The anomalous dimension $\Delta_q^{(\zeta,\nu)}$, if it displays a nonlinear dependence on q , is the signature of multifractal scaling. The index ζ indicates whether the multifractal scaling applies to the bulk ($\zeta = 2$), the one-dimensional boundary ($\zeta = 1$), or to the zero-dimensional boundary (corner) ($\zeta = 0$), provided the plaquette (x, y) is restricted to the corresponding regions of the \mathbb{Z}_2 network model. For $\zeta = 1$ and 0, the index ν distinguishes the case $\nu = O$ when the ζ -dimensional boundary has no edge states in the insulating phase adjacent to the critical point, from the case $\nu = \mathbb{Z}_2$ when the ζ -dimensional boundary has helical edge states in the adjacent insulating phase. We ignore this distinction for multifractal scaling of the bulk wave functions, $\Delta_q^{(2,O)} = \Delta_q^{(2,\mathbb{Z}_2)} = \Delta_q^{(2)}$, since bulk properties are insensitive to boundary effects. We will also consider the case of mixed boundary condition for which we reserve the notation $\nu = \mathbb{Z}_2|O$.

It was shown in [31] that boundary multifractality is related to corner multifractality if it is assumed that conformal invariance holds at the metal-insulator transition in the two-dimensional symplectic universality class. Conversely, the numerical verification of

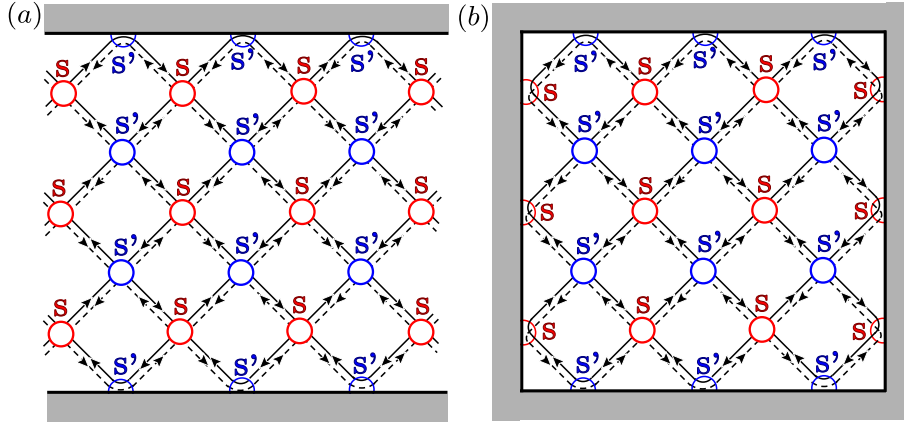


Figure 5. (a) Boundary multifractality is calculated from the wave function amplitudes near a one-dimensional boundary. Periodic (reflecting) boundary conditions are imposed for the horizontal (vertical) boundaries. (b) Corner multifractality is calculated from the wave function amplitudes near a corner with the wedge angle $\vartheta = \pi/2$. Reflecting boundary conditions are imposed along both vertical and horizontal directions. The relationship between the scattering matrix at a node of type S' and the scattering matrix at a node of type S implies that it is a vertical boundary located at nodes of type S that induces an helical edge state when $X > X_l$.

this relationship between boundary and corner multifractality supports the claim that the critical scaling behavior at this metal-insulator transition is conformal. So we want to verify numerically if the consequence of the conformal map $w = z^{\vartheta/\pi}$, namely

$$\Delta_q^{(0,\nu)} = \frac{\pi}{\vartheta} \Delta_q^{(1,\nu)} \quad (120)$$

where ϑ is the wedge angle at the corner, holds. Equivalently, $f^{(\zeta,\nu)}(\alpha)$, which is defined to be the Legendre transformation of $\Delta_q^{(\zeta,\nu)} + dq$, i.e.,

$$\alpha_q^{(\zeta,\nu)} = \frac{d\Delta_q^{(\zeta,\nu)}}{dq} + d, \quad (121)$$

$$f^{(\zeta,\nu)}(\alpha_q) = q\alpha_q^{(\zeta,\nu)} - \Delta_q^{(\zeta,\nu)} - dq + \zeta, \quad (122)$$

must obey

$$\alpha_q^{(0,\nu)} - d = \frac{\pi}{\vartheta} (\alpha_q^{(1,\nu)} - d), \quad (123)$$

$$f^{(0,\nu)}(\alpha) = \frac{\pi}{\vartheta} [f^{(1,\nu)}(\alpha) - 1], \quad (124)$$

if conformal invariance is a property of the metal-insulator transition.

To verify numerically the formulas (120), (123), and (124), we consider the \mathbb{Z}_2 network model with the geometries shown in figure 5. We have calculated wave functions for systems with the linear sizes $L = 50, 80, 120, 150$, and 180 for the two geometries displayed in figure 5. Here, L counts the number of nodes of the same type along a boundary. The number of realizations of the static disorder is 10^5 for each system size.

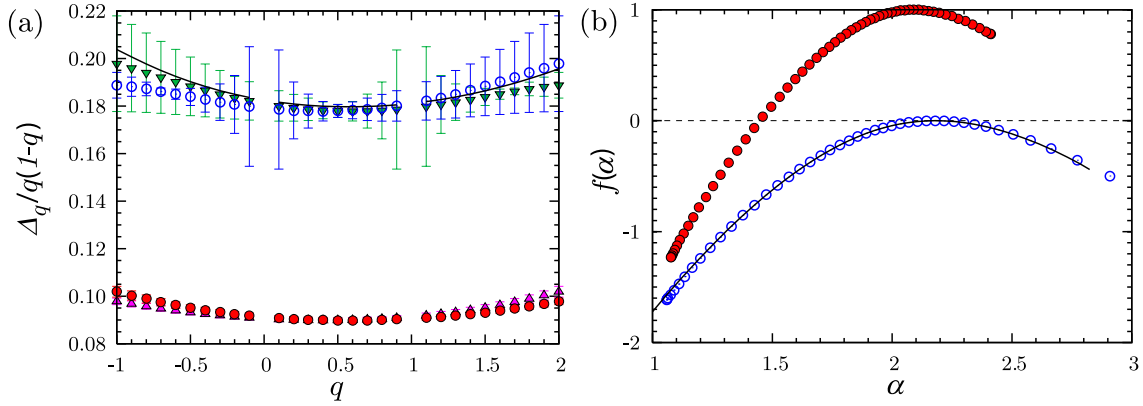


Figure 6. (a) The boundary (filled circles, red) and corner with $\theta = \pi/2$ (open circles, blue) anomalous dimensions at the metal-to- \mathbb{Z}_2 -topological-insulator transition. The solid curve is computed from (120) by using the boundary anomalous dimension as an input. The rescaled Δ_{1-q} confirming the reciprocal relation for boundary and corner multifractality are shown by upper (magenta) and lower (green) triangles, respectively. (b) The multifractal spectra for the boundary (filled circles, red) and the corner (open circles, blue). The solid curve is computed from (123) and (124).

Figure 6(a) shows the boundary anomalous dimensions $\Delta_q^{(1, \mathbb{Z}_2)}$ (filled circles) and the corner anomalous dimensions $\Delta_q^{(0, \mathbb{Z}_2)}$ (open circles). In addition, the anomalous dimensions $\Delta_{1-q}^{(\zeta, \nu)}$ are shown by upper and lower triangles for boundary and corner anomalous dimensions, respectively. They fulfill the reciprocal relation

$$\Delta_q^{(\zeta, \nu)} = \Delta_{1-q}^{(\zeta, \nu)} \quad (125)$$

derived analytically in [34]. Since the triangles and circles are consistent within error bars, our numerical results are reliable, especially between $0 < q < 1$. If we use the numerical values of $\Delta_q^{(1, \mathbb{Z}_2)}$ as inputs in (120) with $\vartheta = \pi/2$, there follows the corner multifractal scaling exponents that are plotted by the solid curve. Since the curve overlaps with the direct numerical computation of $\Delta_q^{(0, \mathbb{Z}_2)}$ within the error bars, we conclude that the relation (120) is valid at the metal-to- \mathbb{Z}_2 -topological-insulator transition.

Figure 6(b) shows the boundary (filled circles) and corner (open circles) multifractal spectra. These multifractal spectra are calculated by using (119), (121), and (122). The numerical values of $\alpha_0^{(\zeta, \mathbb{Z}_2)}$ are

$$\alpha_0^{(1, \mathbb{Z}_2)} = 2.091 \pm 0.002, \quad (126)$$

$$\alpha_0^{(0, \mathbb{Z}_2)} = 2.179 \pm 0.01. \quad (127)$$

The value of $\alpha_0^{(1, \mathbb{Z}_2)}$ is consistent with that reported in [20], while its accuracy is improved. The solid curve obtained from the relations (123) and (124) by using $f^{(1, \mathbb{Z}_2)}(\alpha)$ as an input, coincides with $f^{(0, \mathbb{Z}_2)}(\alpha)$. We conclude that the hypothesis of conformal invariance at the quantum critical point of metal-to- \mathbb{Z}_2 -topological-insulator transition is consistent with our numerical study of multifractal scaling.

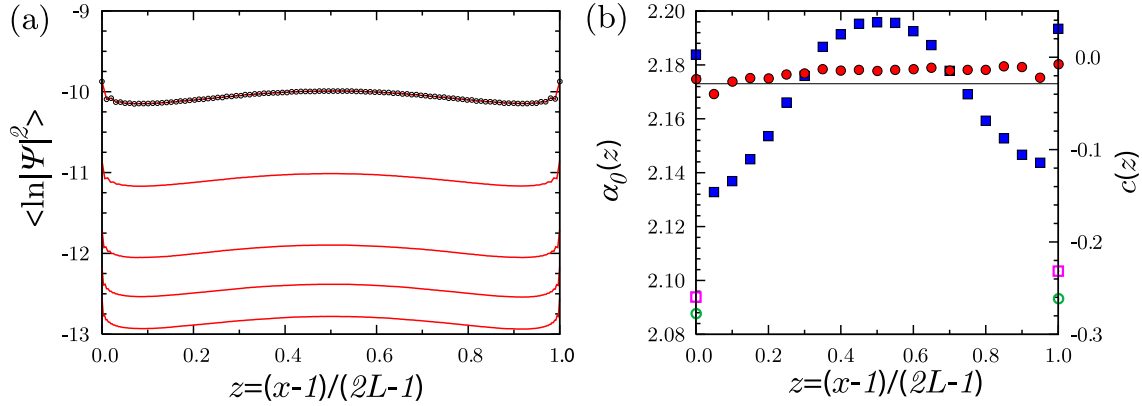


Figure 7. (a) The z dependence of $\langle \ln |\Psi|^2 \rangle_{z,L}$ at the metal-to- \mathbb{Z}_2 -topological-insulator transition in the cylindrical geometry for $L = 50, 80, 120, 150, 180$ from the top to the bottom. (b) The z dependence of $\tilde{\alpha}_0(z)$ (●) and $c(z)$ (■) extrapolated from the system size dependence of $\langle \ln |\Psi|^2 \rangle_{z,L}$ averaged over a small interval of z 's. $\tilde{\alpha}_0(z)$ and $c(z)$ at $z = 0, 1$ without averaging over a small interval of z 's are shown by open circles and squares, respectively. The solid line represents the bulk value of $\alpha_0^{(2)} = 2.173$ computed in [31]. The asymmetry with respect to $z = 0.5$ is due to statistical fluctuations.

At last, we would like to comment on the dependence on z of

$$\langle \ln |\Psi|^2 \rangle_{z,L} \equiv \frac{1}{2L} \sum_{y=1}^{2L} \overline{\ln \left(\sum_{\sigma=\uparrow,\downarrow} |\psi_\sigma(x, y)|^2 \right)} \quad (128)$$

found in [20]. Here, $z \equiv (x-1)/2L$, while x and y denote the positions on the network along its axis and along its circumference, respectively (our choice of periodic boundary conditions imposes a cylindrical geometry). The overline denotes averaging over disorder. Figure 7(a) shows the z dependence of $\langle \ln |\Psi|^2 \rangle_{z,L}$ for different values of L in this cylindrical geometry at the metal-to- \mathbb{Z}_2 -topological-insulator transition. We observe that $\langle \ln |\Psi|^2 \rangle_{z,L}$ becomes a nonmonotonic function of z .

We are going to argue that this nonmonotonic behavior is a finite size effect. We make the scaling ansatz

$$\langle \ln |\Psi|^2 \rangle_{z,L} = -\tilde{\alpha}_0^{(\zeta, \mathbb{Z}_2)}(z) \ln L + c(z), \quad (129)$$

where $\zeta = 1$ if $z = 0, 1$ and $\zeta = 2$ otherwise, while $c(z)$ depends on z but not on L . To check the L dependence of $\langle \ln |\Psi|^2 \rangle_{z,L}$ in figure 7, we average $\langle \ln |\Psi|^2 \rangle_{z,L}$ over a narrow interval of z 's for each L . Figure 7(b) shows the z dependence of $\tilde{\alpha}_0(z)$ (●) and $c(z)$ (■) obtained in this way. In addition, $\tilde{\alpha}_0(z)$ and $c(z)$ calculated for $z = 0, 1$ without averaging over the narrow interval of z 's are shown by open circles and open squares, respectively.

We observe that $\tilde{\alpha}_0(z)$, if calculated by averaging over a finite range of z 's, is almost constant and close to $\alpha_0^{(2)} = 2.173$. In contrast, $\tilde{\alpha}_0(z = 0, 1) \approx 2.09$, if calculated without averaging over a finite range of z 's, is close to $\alpha_0^{(1, \mathbb{Z}_2)} = 2.091$. We also find that $|c(z)|$ increases near the boundaries. We conclude that it is the nonmonotonic dependence of

$|c(z)|$ on z that gives rise to the nonmonotonic dependence of $\langle \ln |\Psi|^2 \rangle_{z,L}$ on z . This finite-size effect is of order $1/\ln L$ and vanishes in the limit $L \rightarrow \infty$.

4.2. Boundary condition changing operator

Next, we impose mixed boundary conditions by either (i) coupling the \mathbb{Z}_2 network model to an external reservoir through point contacts or (ii) by introducing a long-range lead between two nodes from the \mathbb{Z}_2 network model, as shown in figure 8. In this way, when $X > X_l$, a single Kramers' pair of helical edge states indicated by the wavy lines in figure 8 is present on segments of the boundary, while the complementary segments of the boundary are devoid of any helical edge state (the straight lines in figure 8). The helical edge states either escape the \mathbb{Z}_2 network model at the nodes at which leads to a reservoir are attached [the green lines in figure 8(a) and figure 8(b)], or shortcut a segment of the boundary through a nonlocal connection between the two nodes located at the corners [figure 8(c)]. These are the only options that accommodate mixed boundary conditions and are permitted by time-reversal symmetry. As shown by Cardy in [35], mixed boundary conditions are implemented by boundary-condition-changing operators in conformal field theory. Hence, the geometries of figure 8 offer yet another venue to test the hypothesis of two-dimensional conformal invariance at the metal-insulator quantum critical point.

When coupling the \mathbb{Z}_2 network model to an external reservoir, we shall consider two cases shown in figure 8(a) and figure 8(b), respectively.

First, we consider the case of figure 8(a) in which only two nodes from the \mathbb{Z}_2 network model couple to the reservoir. At each of these point contacts, the scattering matrix S that relates incoming to outgoing waves from and to the reservoir is a 2×2 matrix which is invariant under time reversal $s_2 S^* s_2 = S$, and it must be proportional to the unit 2×2 matrix up to an overall (random) phase. Hence, the two-point-contact conductance in the geometry of figure 8(a) is unity, however far the two point contacts are from each other.

Second, we consider the case of figure 8(b) in which there are again two point contacts, however each lead between the \mathbb{Z}_2 network model and the reservoir now supports two instead of one Kramers' doublets. The two point-contact scattering matrices connecting the \mathbb{Z}_2 network model to the reservoirs are now 4×4 matrices, which leads to a non-vanishing probability of backscattering. Hence, this even channel two-point-contact conductance is expected to decay as a function of the separation between the two attachment points of the leads to the network.

To test whether the two-point-contact conductance in figure 8(a) and figure 8(b) do differ as dramatically as anticipated, we have computed numerically the two-point-contact conductance at the quantum critical point $X = X_l$ for a *single realization* of the static disorder. The two-point-contact conductance is calculated by solving for the stationary solution of the time-evolution operator with input and output leads [36]. We choose the cylindrical geometry imposed by periodic boundary conditions along the

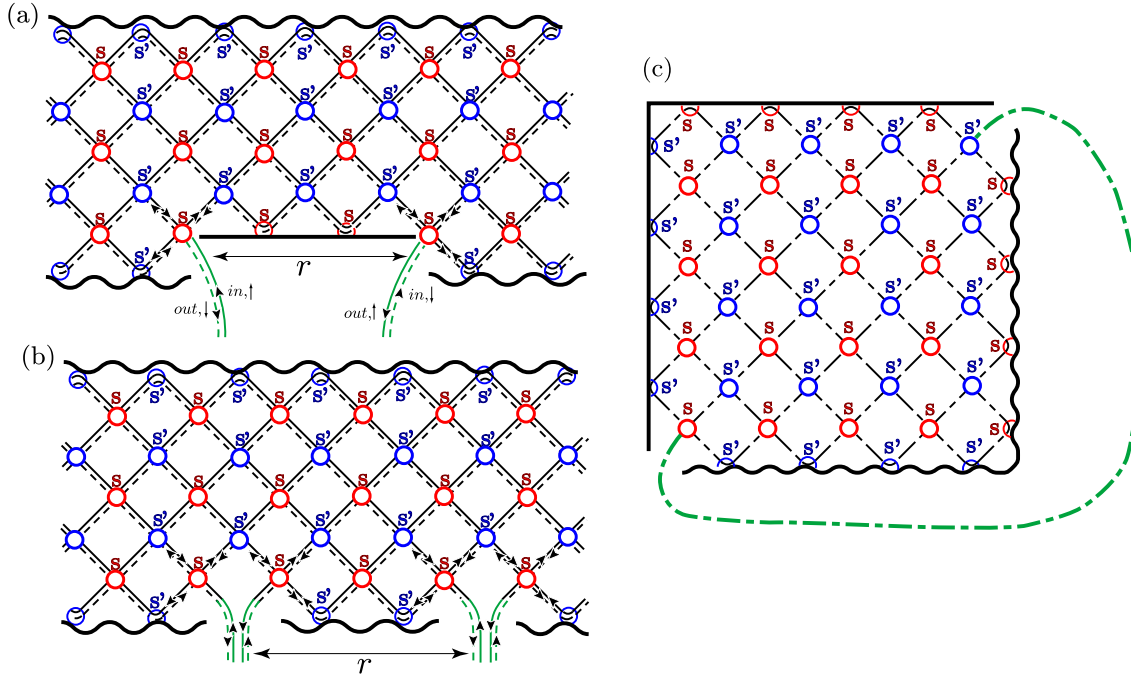


Figure 8. (a) The system with two point contacts (green curves) attached (a) at the two interfaces between different types of boundaries and (b) at the reflecting boundary. The periodic boundary conditions are imposed for the horizontal direction. The thick wavy and solid lines on the edges represent two different types of boundaries, with and without a helical edge mode at $X > X_l$, respectively. (c) Closed network with mixed boundaries. Each dashed line or curve represents a Kramers' doublet.

horizontal directions in figures 8(a) and 8(b) for a squared network with the linear size $L = 200$. Figure 9(a) shows with the symbol \bullet the dependence on r , the distance between the two contacts in figure 8(a), of the dimensionless two-point-contact conductance g . It is evidently r independent and unity, as expected. Figure 9(a) also shows with the symbol \circ the dependence on r of the dimensionless two-point conductance g for leads supporting two Kramers' doublet as depicted in figure 8(b). Although it is not possible to establish a monotonous decay of the two-point-contact conductance for a single realization of the static disorder, its strong fluctuations as r is varied are consistent with this claim.

We turn our attention to the closed geometry shown in figure 8(c). We recall that it is expected on general grounds that the moments of the two-point conductance in a network model at criticality, when the point contacts are far apart, decay as power laws with scaling exponents proportional to the scaling exponents $\Delta_q^{(\zeta, \nu)}$ [36, 37]. Consequently, after tuning the \mathbb{Z}_2 network model to criticality, the anomalous dimensions at the node (corner) where the boundary condition is changed must vanish,

$$\Delta_q^{(0, \mathbb{Z}_2 | 0)} = 0, \quad (130)$$

since the two-point-contact conductance in figure 8(a) is r independent. (130) is another signature of the nontrivial topological nature of the insulating side at the Anderson

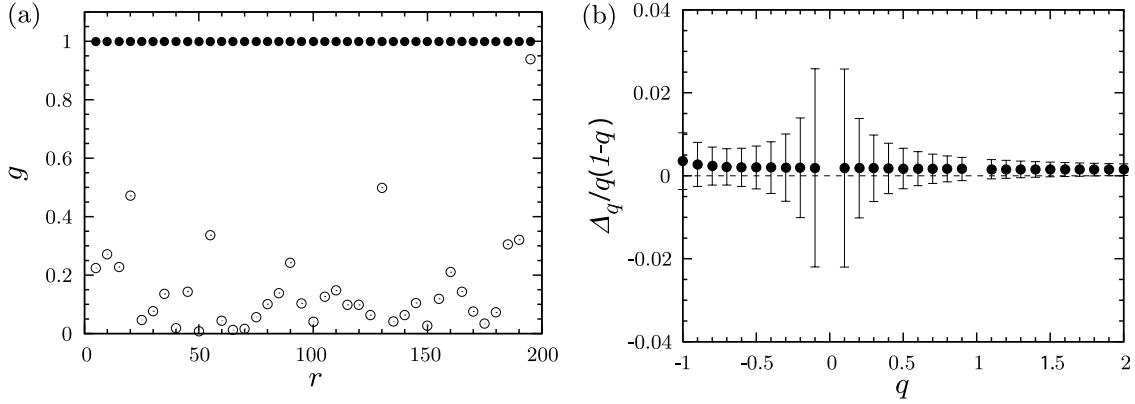


Figure 9. (a) The distance r dependence of the two-point-contact conductance g for the mixed boundary (\bullet) and the reflecting boundary (\circ). (b) The zero dimensional anomalous dimension $\Delta_q^{(0, \mathbb{Z}_2 | O)}$ obtained from the wave function amplitude on the link which connecting the boundary condition changing points.

transition that we want to test numerically. Thus, we consider the geometry figure 8(c) and compute numerically the corner anomalous dimensions. This is done using the amplitudes of the stationary wave function restricted to the links connecting the two corners where the boundary conditions are changed. Figure 9(b) shows the numerical value of the corner anomalous dimension $\Delta_q^{(0, \mathbb{Z}_2 | O)}$. The linear sizes of the network are $L = 50, 80, 120, 150$, and 180 and the number of disorder realizations is 10^5 for each L . We observe that $\Delta_q^{(0, \mathbb{Z}_2 | O)}$ is zero within the error bars, thereby confirming the validity of the prediction (130) at the metal-to- \mathbb{Z}_2 -topological-insulator transition.

5. Conclusions

In summary, we have mapped the \mathbb{Z}_2 network model to a 4×4 Dirac Hamiltonian. In the clean limit of this Dirac Hamiltonian, we expressed the Kane-Mele \mathbb{Z}_2 invariant as an $SU(2)$ Wilson loop and computed it explicitly. In the presence of weak time-reversal symmetric disorder, the NLSM that can be derived out of this Dirac Hamiltonian describes the metal-insulator transition in the \mathbb{Z}_2 network model and yields bulk scaling exponents that belong to the standard two-dimensional symplectic universality class; an expectation confirmed by the numerics in [19] and [20]. A sensitivity to the \mathbb{Z}_2 topological nature of the insulating state can only be found by probing the boundaries, which we did numerically in the \mathbb{Z}_2 network model by improving the quality of the numerical study of the boundary multifractality in the \mathbb{Z}_2 network model.

References

- [1] Roland Winkler 2003 “*Spin-orbit coupling effects in two-dimensional electron and hole systems*,” (Springer-Verlag Berlin Heidelberg)
- [2] Hikami S, Larkin A I and Nagaoka Y 1980 *Prog. Theor. Phys.* **63** 707

- [3] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 226801
- [4] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 146802
- [5] Bernevig B A and Zhang S C 2006 *Phys. Rev. Lett.* **96** 106802
- [6] Bernevig B A, Hughes T L and Zhang S C 2006 *Science* **314** 1757
- [7] König M, Wiedmann S, Brüne C, Roth A, Buhmann H, Molenkamp L W, Qi X L and Zhang S C 2007 *Science* **318** 766
- [8] Moore J E and Balents L 2007 *Phys. Rev. B* **75** 121306(R)
- [9] Roy R 2009 *Phys. Rev. B* **79** 195322
- [10] Fu L, Kane C L and Mele E J 2007 *Phys. Rev. Lett.* **98** 106803
- [11] Hsieh D, Qian D, Wray L, Xia Y, Hor Y, Cava R and Hasan M Z 2008 *Nature* **452** 970
- [12] Hsieh D, Xia Y, Wray L, Qian D, Pal A, Dil J H, Osterwalder J, Meier R, Bihknayer G, Kane C L, Hor Y, Cava R, and Hasan M 2009 *Science* **323** 919
- [13] Xia Y, Qian D, Hsieh D, Wray L, Pal A, Lin H, Bansil A, Grauer D, Hor Y S, Cava R J, and Hasan M Z 2009 *Nature Phys.* **5** 398
- [14] Hsieh D, Xia Y, Qian D, Wray L, Dil J H, Meier F, Osterwalder J, Patthey L, Checkelsky J G, Ong N P, Fedorov A V, Lin H, Bansil A, Grauer D, Hor Y S, Cava R J, and Hasan M Z 2009 *Nature* **460** 1101
- [15] Chen Y L, Analytis J G, Chu J-H, Liu Z K, Mo S-K, Qi X L, Zhang H J, Lu D H, Dai X, Fang Z, Zhang S C, Fisher I R, Hussain Z, and Shen Z-X 2009 *Science* **325** 178
- [16] Thouless D J, Kohmoto M, Nightingale M P and den Nijs M 1982 *Phys. Rev. Lett.* **49** 405
- [17] Fu L and Kane C L 2006 *Phys. Rev. B* **74** 195312
- [18] Onoda M, Avishai Y and Nagaosa N 2007 *Phys. Rev. Lett.* **98** 076802
- [19] Obuse H, Furusaki A, Ryu S and Mudry C 2007 *Phys. Rev. B* **76** 075301
- [20] Obuse H, Furusaki A, Ryu S and Mudry C 2008 *Phys. Rev. B* **78** 115301
- [21] Chalker J T and Coddington P D 1988 *J. Phys. C* **21** 2665
- [22] Kramer B, Ohtsuki T and Kettemann S 2005 *Phys. Rep.* **417** 211
- [23] Wegner F J 1979 *Z. Phys. B* **35** 207
- [24] Fendley P 2001 *Phys. Rev. B* **63** 104429
- [25] Ryu S, Mudry S, Obuse H and Furusaki A 2007 *Phys. Rev. Lett.* **99** 116601
- [26] Ostrovsky P M, Gornyi I V and Mirlin A D 2007 *Phys. Rev. Lett.* **98** 256801
- [27] Bardarson J H, Tworzydło J, Brouwer P W, and Beenakker C W J, *Phys. Rev. Lett.* **99**, 106801 (2007).
- [28] Nomura K, Koshino M, and Ryu S 2007 *Phys. Rev. Lett.* **99** 146806
- [29] Schnyder A P, Ryu S, Furusaki A and Ludwig A W W 2008 *Phys. Rev. B* **78** 195125
- [30] Subramaniam A R, Gruzberg I A, Ludwig A W W, Evers F, Mildenerger A and Mirlin A D 2006 *Phys. Rev. Lett.* **96** 126802
- [31] Obuse H, Subramaniam A R, Furusaki A, Gruzberg I A and Ludwig A W W 2007 *Phys. Rev. Lett.* **98** 156802
- [32] Ho C M and Chalker J T 1996 *Phys. Rev. B* **54** 8708
- [33] Ludwig A W W, Fisher M P A, Shankar R and Grinstein G 1994 *Phys. Rev. B* **50** 7526
- [34] Mirlin A D, Fyodorov Y V, Mildenerger A and Evers F 2006 *Phys. Rev. Lett.* **97** 046803
- [35] Cardy J L 1989 *Nucl. Phys. B* **324** 581
- [36] Janssen M, Metzler M and Zirnbauer M R 1999 *Phys. Rev. B* **59** 15836
- [37] Klesse R and Zirnbauer M R 2001 *Phys. Rev. Lett.* **86** 2094